Weak Solvability Theorem for Forward-Backward SDEs

François Delarue, Giuseppina Guatteri

Weak Existence and Uniqueness for Forward-Backward SDEs

François Delarue[(a),*] and Giuseppina Guatteri[(b),†]

(a) Laboratoire de Probabilités et Modèles Aléatoires,
Université Paris VII, UFR de Mathématiques, Case 7012,
2, Place Jussieu, 75251 Paris Cedex 05 - FRANCE.
(b) Dipartimento di Matematica, "Francesco Brioschi", Politecnico di Milano,
Piazza Leonardo da Vinci, 32, 20 133 Milano - ITALY.

29th December 2005

Abstract

We aim to establish the existence and uniqueness of weak solutions to a suitable class of non-degenerate deterministic FBSDEs with a one-dimensional backward component. The classical Lipschitz framework is partially weakened: the diffusion matrix and the final condition are assumed to be space Hölder continuous whereas the drift and the backward driver may be discontinuous in x. The growth of the backward driver is allowed to be at most quadratic with respect to the gradient term.

The strategy holds in three different steps. We first build a well controlled solution to the associated PDE and as a bypass product a weak solution to the forward-backward system. We then adapt the "decoupling strategy" introduced in the *four step scheme* of Ma, Protter and Yong [30] to prove uniqueness.

AMS 2000 Subject Classification. Primary: 65C30. Secondary: 35K55, 60H10, 60H30.

Keywords. FBSDEs, Gradient Estimates, Quasi-linear PDEs, Calderón and Zygmund Estimates, Schauder's Estimates, Weak Existence and Uniqueness.

1 Introduction

General Setting. Forward Backward SDEs were introduced in 1993 by Antonelli [1] as an extension of the earlier theory of Backward SDEs due to Pardoux and Peng [32] and [33]. Such equations strongly couple a stochastic differential equation to a backward one: the coefficients of each component explicitly depend on the solution of the other one. In a rough way, the resulting system writes as a kind of stochastic two-point boundary value problem:

$$(E) \quad \begin{cases} \forall t \in [0,T], \\ X_t = x + \int_0^t b(s,X_s,Y_s,Z_s)ds + \int_0^t \sigma(s,X_s,Y_s)dB_s, \\ Y_t = G(X_T) + \int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s^* \sigma(s,X_s,Y_s)dB_s. \end{cases}$$

*delarue@math.jussieu.fr (corresponding author)
†guatteri@mate.polimi.it

The whole paper then focuses on the solvability of (E). For this reason, we do not discuss in detail the application fields of the FBSDE theory and just refer to the monograph of Ma and Yong [29] for typical examples arising in mathematical finance or in control problems.

Existing Literature. Due to the strong coupling between both components of (E), it is well understood that solving a forward-backward problem requires much more effort than solving a SDE of Itô or backward type. In particular, the strategy based on Picard's fixed point theorem is not so successfull as in the so-called decoupled setting considered by Pardoux and Peng [33] (*i.e.* $b = b(t,x)$ and $\sigma = \sigma(t,x)$). Applying this method, Antonelli [1] establishes the unique solvability of Lipschitz continuous FBSDEs defined on intervals of small length: relevant counter examples in [1] show that both existence and uniqueness may fail in this frame for arbitrarily prescribed time duration T.

During the last ten years, many papers have exhibited sufficient conditions to ensure the unique solvability on an interval of arbitrary length. Generally speaking, two families of methods have been considered.

The first one applies under monotonicity assumptions to deterministic and stochastic coefficients. Different types of conditions have been investigated in this framework and we refer to Hu and Peng [17], Peng and Wu [35], and Yong [41] on the one hand and to Pardoux and Tang [34] on the other hand for a precise review of the most common hypotheses in this setting.

The second approach relies on the connection between SDEs with deterministic coefficients and non-linear PDEs. It is now well-known that a deterministic FBSDE of type (E) provides a probabilistic representation of the solutions of a system of quasi-linear PDEs: this explains why FBSDEs are usually described as extensions of the Feynman-Kac formula. In the current paper, the backward component of (E) is assumed to be one-dimensional and the underlying system of PDEs reduces to a PDE of the following form (with $a(t,x) = (\sigma\sigma^*)(t,x)$):

$$(\mathcal{E}) \quad \begin{cases} \partial_t u(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,x,u(t,x)\big)\partial^2_{x_i,x_j} u(t,x) \\ \quad + \displaystyle\sum_{i=1}^{d} b_i\big(t,x,u(t,x),\nabla_x u(t,x)\big)\partial_{x_i} u(t,x) + f\big(t,x,u(t,x),\nabla_x u(t,x)\big) = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d, \\ u(T,x) = G(x), \ x \in \mathbb{R}^d. \end{cases}$$

This deep connection permits to apply the huge literature devoted to non-linear PDEs to investigate the theory of forward-backward equations. For instance, referring to the famous monograph of Ladyzhenskaya *et al.* [26], Ma, Protter and Yong [30] establish, for smooth coefficients b, f, σ and G, the strong solvability of non-degenerate deterministic FBSDEs: the diffusion matrix a is assumed to be uniformly elliptic to overcome the inherent strong coupling. This approach, known as the *four step scheme*, is probably the most popular existing one on the topic. In Delarue [7], the first author relaxes the regularity assumption required in Ma *et al.* [30] by combining the short time theory of Antonelli [1] to *a priori* estimates of the gradient of the solutions of $(\mathcal{E})$. These gradient estimates are given in Ladyzhenskaya *et al.* [26], Chapter VII, Section 6, and proved with stochastic arguments in Delarue [8].

Objective of the Paper. In the whole paper, following Ma *et al.* [30] and Delarue [7], the coefficients of (E) are assumed to be deterministic and the diffusion matrix to be non-degenerate. As already said, the backward component is also chosen to be one-dimensional. Here is the novelty compared to the previous references: the matrix a is space Hölder continuous, uniformly in t (a is smooth in [30] and Lipschitz continuous in x in [7]), the coefficients b and f may be discontinuous in time and space (b and f are smooth in [30] and b is monotonous and continuous in x and f is Lipschitz in x in [7]), the final condition G is Hölder continuous (it belongs to $C^{2+\alpha}(\mathbb{R})$ in [30] and is Lipschitz in [7]), and finally, thanks to the one-dimensional assumption, the growth of f is at most quadratic in Z (it is at most linear in [30] and [7]). In this setting, we establish the existence and uniqueness of a so-called "weak solution" to the stochastic system (E), and as a bypass product the unique solvability of the PDE $(\mathcal{E})$ (cf Theorem 2.1). Since a is just Hölder continuous in x, the strong solvability of (E) may fail.

The notions of weak existence and uniqueness for forward-backward equations are very similar to the ones considered for classical SDEs. Referring to the basic definitions given in Rogers and Williams [36], Chapter V, Section 3, the reader can guess without much effort that the word "weak" indicates that existence does not hold on an arbitrarily prescribed Brownian set-up and that uniqueness just holds in law. For a complete overview on weak solutions to FBSDEs, we refer to the paper of Antonelli and Ma [2]. The reader can also find another example of weak existence in Lejay [27]. However, this latter result applies to specific coefficients deriving from a divergence form operator and no uniqueness property is established in this case.

We then feel that our paper is somehow the first to draw up a clear frame for which both existence and uniqueness hold in the weak sense.

Strategy. Our strategy aims to adapt the skeleton of the *four step scheme* of Ma *et al.* [30] to the weak point of view. Build first a solution u to the PDE $(\mathcal{E})$ from a regularization procedure and deduce the weak solvability of (E) from the theory of Stroock and Varadhan [37]. Apply then Itô's formula to u to break the strong relationship between the forward and backward components and derive the uniqueness of the distribution of the solution. This approach thus turns out to be a "decoupling strategy". To handle in this frame the quadratic growth of the coefficient f, we successfully apply the ideas developed by Kobylanski [20] in the quadratic decoupled backward case.

The whole difficulty consists in fact in controlling the derivatives of u: to apply efficiently the Itô formula to u, the partial derivatives of u of order one in t and of order one and two in x must be estimated in a relevant way. This procedure is far from being simple in our poor setting, and at the opposite of the existing literature, the partial derivatives of u of order one in t and of order two in x are just controlled in our frame in suitable L^p spaces. The main argument to establish these bounds follows from the Calderón and Zygmund theory.

Mention finally that we have tried to detail the proofs of most of the controls used in the paper and to avoid as much as possible to refer the reader to too many different existing estimates.

Organization of the Paper. We first detail in Section 2 general assumption and notation and remind the reader of the notion of weak uniqueness. We also specify the statement of the main result. In Section 3, we give crucial *a priori* estimates of the solution and of its derivatives in the smooth framework. These estimates permit to establish in Section 4 the existence of a "well controlled solution" to $(\mathcal{E})$: we then derive the weak unique solvability of the forward-backward equation. In Sections 5, 6 and 7, we prove the previous *a priori* estimates: Section 5 gives a general overview of the strategy. As a conclusion, we discuss in Section 8 the strong solvability of (E) and give further interests of our results.

2 Assumption and Notation

In this section, we first detail the assumptions on the coefficients b, f, σ and G. We also recall the classical definition of strong solutions to forward-backward equations and detail, in this framework, the connection with quasi-linear PDEs. We then investigate the notion of weak solutions and state the main result of the paper. We finally discuss the strategy of the proof.

In the whole paper, the Euclidean norm on $\mathbb{R}^n$, $n \geq 1$, is denoted by $|\cdot|$, and the associated scalar product by $\langle \cdot, \cdot \rangle$. The n-uple $(e_1, \dots, e_n)$ then denotes the canonical basis of $\mathbb{R}^n$, and $B(x_0, \rho)$ (resp. $\bar{B}(x_0, \rho)$), $x_0 \in \mathbb{R}^n, \rho > 0$, the open (resp. closed) Euclidean ball of center x_0 and of radius ρ.

2.1 Coefficients of the Equation

For given $d \in \mathbb{N}^*$ and $T \in \mathbb{R}_+^*$, we consider the following Borel-measurable coefficients:

$$b : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d,\ f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R},\ \sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d \times d},\ G : \mathbb{R}^d \to \mathbb{R}.$$

Assumption (A) We say that the former functions b, f, σ and G satisfy Assumption **(A)** if there exist five constants $\alpha_0 > 0$, H, K, $\lambda > 0$ and Λ, such that:

(A.1) $\forall t \in [0,T],\ \forall (x,y,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$,

$$\begin{aligned} &|(b, \sigma, G)(t,x,y,z)| \leq \Lambda\big(1 + |y| + |z|\big), \\ &|f(t,x,y,z)| \leq \Lambda\big(1 + |y| + |z|^2\big). \end{aligned}$$

(A.2) $\forall (t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R},\ \forall \zeta \in \mathbb{R}^d,\ \langle \zeta, a(t,x,y)\zeta \rangle \geq \lambda |\zeta|^2$, where $a(t,x,y) = \sigma\sigma^*(t,x,y)$.

(A.3) $\forall t \in [0,T],\ \forall (x,y,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d,\ \forall (y', z') \in \mathbb{R} \times \mathbb{R}^d$:

$$\begin{aligned} &|a(t,x,y) - a(t,x,y')| \leq K\,|y - y'|, \\ &|b(t,x,y,z) - b(t,x,y',z')| \leq K\big(|y - y'| + |z - z'|\big), \\ &|f(t,x,y,z) - f(t,x,y',z')| \leq K\big(1 + |z| + |z'|\big)\big(|y - y'| + |z - z'|\big). \end{aligned}$$

(A.4) $\forall t \in [0,T],\ \forall (x,x',y) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$:

$$|a(t,x',y) - a(t,x,y)| + |G(x') - G(x)| \leq H\,|x' - x|^{\alpha_0}.$$

Assumption **(A.3)** for f can be written in a more tractable way:

(A.5) $\forall t \in [0,T],\ \forall x \in \mathbb{R}^d,\ \forall (y,z),\ (y',z') \in \mathbb{R} \times \mathbb{R}^d$:

$$|f(t,x,y,z) - f(t,x,y',z')| \leq K\big(1 + 2|z| + |z' - z|\big)\big(|y - y'| + |z - z'|\big).$$

2.2 Strong Solutions to Forward-Backward Equations

Recall now several properties of strong solutions to forward-backward systems. Consider to this end a filtered probability space $(\Omega, \{\mathcal{F}_s\}_{0 \leq s \leq T}, \mathbb{P})$ satisfying the usual conditions and endowed with an $\{\mathcal{F}_s\}_{0 \leq s \leq T}$-Brownian motion $(B_s)_{0 \leq s \leq T}$ with values in $\mathbb{R}^d$. To the coefficients (b, f, σ, G) and to a given initial condition $(t,x) \in [0,T] \times \mathbb{R}^d$, we associate the following couple of stochastic differential equations:

$$(E)\ \begin{cases} \forall s \in [t,T], \\ X_s = x + \displaystyle\int_t^s b(r, X_r, Y_r, Z_r)dr + \int_t^s \sigma(r, X_r, Y_r)dB_r, \\ Y_s = G(X_T) + \displaystyle\int_s^T f(r, X_r, Y_r, Z_r)dr - \int_s^T \langle Z_r, \sigma(r, X_r, Y_r)dB_r \rangle. \end{cases}$$

Specify now the sense given to the solution (X,Y,Z). Introduce to this end, for $t \in [0,T]$ and $q \geq 1$, the following spaces:

$$
\begin{cases}
H^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^q): & \text{space of } \{\mathcal{F}_s\}_{t \leq s \leq T} - \text{progressively measurable processes} \\
& v : \Omega \times [t,T] \to \mathbb{R}^q \mid \|v\|_2^2 \equiv \mathbb{E}[\int_t^T |v_s|^2 \, ds] < +\infty, \\
& \\
S^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^q): & \text{space of continuous } \{\mathcal{F}_s\}_{t \leq s \leq T} - \text{adapted processes} \\
& v : \Omega \times [t,T] \to \mathbb{R}^q \mid \|v\|_*^2 \equiv \mathbb{E}[\sup_{s \in [t,T]} |v_s|^2] < +\infty.
\end{cases}
$$

A triple (X,Y,Z) is then said to be a strong solution to the FBSDE (E) with initial condition (t,x) if:

1. $X \in S^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^d)$, $Y \in S^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R})$, $Z \in H^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^d)$,
2. $\mathbb{P}$ almost-surely, (X,Y,Z) satifies (E).

Recall from Delarue [7] that there exists a unique strong solution to (E) if the coefficients (b,f,σ,G) are bounded in (t,x) and at most linear in (y,z), Lipschitz continuous in (x,y,z) uniformly in t, and if σ is continuous and satisfies the ellipticity condition **(A.2)**. The solution is usually denoted by $(X^{t,x}, Y^{t,x}, Z^{t,x})$: the superscript (t,x) denotes the initial condition of the diffusion X.

Moreover, according to Ladyzhenskaya *et al.* [26], Chapter VII, Theorem 7.1, and to Ma, Protter, Yong [30], if the coefficients (b,f,σ,G) are smooth, *i.e.* infinitely differentiable with respect to the variables t,x,y and z, and bounded, with bounded derivatives of any order, then the following quasi-linear PDE:

$$
(\mathcal{E}) \begin{cases}
\partial_t u(t,x) + \dfrac{1}{2} \displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,x,u(t,x)\big) \partial^2_{x_i,x_j} u(t,x) \\
\quad + \displaystyle\sum_{i=1}^{d} b_i\big(t,x,u(t,x),\nabla_x u(t,x)\big) \partial_{x_i} u(t,x) + f\big(t,x,u(t,x),\nabla_x u(t,x)\big) = 0, \ (t,x) \in [0,T[\times \mathbb{R}^d, \\
u(T,x) = G(x), \ x \in \mathbb{R}^d,
\end{cases}
$$

admits a unique bounded solution with a bounded gradient in the space $\mathcal{C}^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$. Moreover, the gradient is Hölder continuous on $[0,T] \times \mathbb{R}^d$ and the derivatives of order one in t and of order two in x are also bounded and Hölder continuous on $[0,T] \times \mathbb{R}^d$. In such a case, the following connection holds between u and $(X^{t,x}, Y^{t,x}, Z^{t,x})$:

$$
\forall s \in [t,T], \ Y^{t,x}_s = u(s, X^{t,x}_s), \ Z^{t,x}_s = \nabla_x u(s, X^{t,x}_s). \tag{2.1}
$$

Conversely, the solution u writes:

$$
\forall (t,x) \in [0,T] \times \mathbb{R}^d, \ u(t,x) = Y^{t,x}_t. \tag{2.2}
$$

Note that several papers have extended the connection between forward-backward equations and quasi-linear PDEs to other kinds of solutions: Pardoux and Tang [34] consider viscosity solutions and Delarue [9] focuses on the Sobolev sense.

2.3 Weak Solutions and Main Result

Earlier results of existence and uniqueness, e.g. Ma, Protter, Yong [30], Pardoux and Tang [34] or Delarue [7], do not apply under Assumption **(A)**. First, the growth of the driver f is quadratic in z, and second, the coefficients b, f, σ and G are not Lipschitz in x (the coefficients b and f may even be discontinuous with respect to the space variable).

Review now the consequences of each of these points on the solvability of (E).

Focus on the growth of f, and recall that the paper of Kobylanski [20] investigates the existence and uniqueness of solutions to backward SDEs with quadratic drivers. Generally speaking, there is a double price to pay to allow the coefficient f to be quadratic. First, the process Y has to live in the one-dimensional real space: this is the case in our setting. Second, uniqueness of solutions to quadratic backward equations just holds for processes Y with uniformly bounded trajectories.

Moreover, the rather "weak" regularity properties of b, f, σ and G make the classical framework of FBSDEs unsuitable. This is well understood since the strong solvability of SDEs with Hölder continuous and non-degenerate diffusion coefficients may fail: see Barlow [3] for a one-dimensional counter-example. At the opposite, the so-called "weak theory" seems particulary relevant in our setting: the point of view of Stroock and Varadhan [37] may apply since the diffusion matrix a is uniformly elliptic. We then seek in the sequel for a weak solution to the forward-backward system (E).

Note again that Antonelli and Ma [2] as well as Lejay [27] already introduced this concept. For the sake of completeness, we remind the reader of the basic notions and first define the framework of any weak theory for SDEs:

Definition 2.1 *A four-uple $(\Omega, \{\mathcal{F}\}, \mathbb{P}, B)$ is said to be a standard set-up if:*

1. $(\Omega, \{\mathcal{F}_s\}_{0\le s\le T}, \mathbb{P})$ *is a filtered probability space satisfying the* usual conditions*,*
2. $(B_s)_{0\le s\le T}$ *is an $\mathbb{R}^d$-valued Brownian motion on the above space.*

According to our previous discussion on the boundedness of the solutions to quadratic BSDEs, we introduce, for a standard set-up $(\Omega, \{\mathcal{F}\}, \mathbb{P}, B)$, a real $t \in [0, T]$ and an integer $q \ge 1$, the following class of processes:

$$\left\{ \begin{array}{ll} S^\infty_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^q): & \text{space of continuous } \{\mathcal{F}_s\}_{t\le s\le T}-\text{ adapted processes} \\ & v : \Omega \times [t, T] \to \mathbb{R}^q \mid \|v\|_\infty \equiv \operatorname{essup}_{\omega\in\Omega} \sup_{s\in[t,T]} |v_s| < +\infty. \end{array} \right.$$

We are now in position to give the definition of a weak solution to the forward-backward system (E):

Definition 2.2 *For $(t, x) \in [0, T] \times \mathbb{R}^d$, a triple of processes (X, Y, Z) is said to be a weak solution of (E) with initial condition (t, x) if there exists a standard set-up $(\Omega, \{\mathcal{F}\}, \mathbb{P}, B)$ such that:*

1. $X \in S^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^d)$, $Y \in S^\infty_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R})$, $Z \in H^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^d)$,
2. $\mathbb{P}$ *almost-surely, (X, Y, Z) satisfies (E).*

Of course, any strong solution gives rise to a weak solution. In short, strong existence implies weak existence. Note also that the same holds for uniqueness: strong uniqueness implies weak uniqueness, see e.g Antonelli and Ma [2] or Delarue [7], Remark 1.6.

Here is the result that we establish in this article:

Theorem 2.1 *Let $(t, x) \in [0, T] \times \mathbb{R}^d$. Then, under Assumption* **(A)***, the Forward-Backward SDE (E) admits a weak solution $((\Omega, \{\mathcal{F}\}, \mathbb{P}, B), (X, Y, Z))$ with initial condition (t, x).*

Moreover, if $((\tilde{\Omega}, \{\tilde{\mathcal{F}}\}, \tilde{\mathbb{P}}, \tilde{B}), (\tilde{X}, \tilde{Y}, \tilde{Z}))$ denotes another weak solution with initial condition (t, x), *then the distributions* $(\tilde{B}, \tilde{X}, \tilde{Y}, \tilde{Z})(\tilde{\mathbb{P}})$ *and* $(B, X, Y, Z)(\mathbb{P})$ *on the space* $\mathcal{C}([t,T], \mathbb{R}^d) \times \mathcal{C}([t,T], \mathbb{R}^d) \times \mathcal{C}([t,T], \mathbb{R}) \times L^2([t,T], \mathbb{R}^d)$ *are equal.*

From an analytical point of view, there exists a unique solution to the PDE $(\mathcal{E})$ *in the space :*

$$
\begin{aligned}
&\mathcal{V} \equiv \big\{ u \in \mathcal{C}^0([0,T] \times \mathbb{R}^d, \mathbb{R}) \cap \mathcal{C}^{0,1}([0,T[\times \mathbb{R}^d, \mathbb{R}) \cap W^{1,2,d+1}_{\mathrm{loc}}([0,T[\times \mathbb{R}^d, \mathbb{R}), \\
&\quad \exists\, \gamma > 0, \sup_{(t,x)\in[0,T[\times\mathbb{R}^d} \big(|u(t,x)| + (T-t)^{1/2-\gamma} |\nabla_x u(t,x)| \big) < +\infty \big\}, \\
&\text{with } W^{1,2,d+1}_{\mathrm{loc}}([0,T[\times \mathbb{R}^d, \mathbb{R}) \\
&\quad \equiv \big\{ u : [0,T[\times \mathbb{R}^d \to \mathbb{R},\ |u|,\ |\nabla_x u|,\ |\nabla^2_{x,x} u|,\ |\partial_t u| \in L^{d+1}_{\mathrm{loc}}([0,T[\times \mathbb{R}^d, \mathbb{R}) \big\}.
\end{aligned}
$$

The process (Y,Z) *can then be chosen to satisfy:*

$$
\forall s \in [t,T],\ Y_s = u(s, X_s),\ \forall s \in [t,T[,\ Z_s = \nabla_x u(s, X_s).
$$

2.4 Strategy of the Proof

Say now a word about the strategy used to establish Theorem 2.1.

Existence. Start first with existence of a weak solution. Generally speaking, the method is rather simple. Build first a solution $u \in \mathcal{V}$ to the PDE $(\mathcal{E})$ and solve in a weak sense the following SDE:

$$
\forall s \in [t,T],\ dX_s = b(s, X_s, u(s,X_s), \nabla_x u(s,X_s))ds + \sigma(s, X_s, u(s,X_s))dB_s. \tag{2.3}
$$

Thanks to the Itô formula (or Itô-Krylov in our frame since u admits generalized derivatives of order one in t and order two in x), deduce then that the couple $(Y_s, Z_s)_{t\le s\le T} \equiv (u(s,X_s), \nabla_x u(s,X_s))_{t\le s\le T}$ satisfies the required backward equation on the standard set-up given by the forward component X.

The whole difficulty is then hidden in the construction of the solution u. A classical strategy to investigate the solvability of the PDE $(\mathcal{E})$ consists in deriving the existence of a solution through compactness arguments. For example, for mollifiers $(b_n, f_n, \sigma_n, G_n)_{n\ge1}$, find uniform *a priori* Hölder controls of the associated solutions $(u_n)_{n\ge1}$ (that exist in the regularized framework) and of their partial derivatives $(\nabla_x u_n)_{n\ge1}$, $(\nabla^2_{x,x} u_n)_{n\ge1}$ and $(\partial_t u_n)_{n\ge1}$ in terms of known parameters and extract a converging subsequence from the Arzelà-Ascoli theorem. Such a method holds essentially for Hölder continuous coefficients (b, f, σ, G) for which the Schauder theory applies. In our frame, since b and f may be discontinuous in t and x, we are just able to establish similar Hölder controls for $(u_n)_{n\ge1}$ and $(\nabla_x u_n)_{n\ge1}$ and to prove in addition from the Calderón-Zygmund point of view that $(\nabla^2_{x,x} u_n)_{n\ge1}$ and $(\partial_t u_n)_{n\ge1}$ are bounded in suitable L^p spaces. This permits to extract a subsequence for which the second derivatives in x and the first derivative in t converge weakly in L^p.

Note nevertheless that this strategy is not the only one. For example, in Guatteri and Lunardi [16], the authors derive directly the smoothing property of the solution u to $(\mathcal{E})$ from a fixed point argument performed in a suitable topological space. The key tool to achieve the strategy is the regularizing property of the evolution operator associated to a linearized version of the PDE $(\mathcal{E})$ (see Lunardi [28]). However, the fixed point procedure requires stronger regularity properties on the coefficients: (b, f, σ) are uniformly Lipschitz continuous in the first variable and twice differentiable in the other variables with uniformly bounded second order derivatives in x, and as in **(A)**, f is quadratic in z, and the final condition G is in $\mathcal{C}^\beta(\mathbb{R}^d)$, $\beta > 0$.

Uniqueness. The proof of weak uniqueness relies on a non-trivial variation of the uniqueness property given in the *four-step scheme* of Ma, Protter and Yong [30]. To illustrate this approach, focus first on the strong uniqueness framework and assume that X, given by (2.3), is a strong solution. As explained above, the triple $(X, Y = u(\cdot, X), Z = \nabla_x u(\cdot, X))$ satisfies the forward-backward equation. Denote now by (U, V, W) another solution to the FBSDE (E) with the same initial condition. Instead of studying the difference $X - U$ and $(Y - V, Z - W)$ as done in Delarue [7], the strategy introduced by Ma, Protter and Yong [30] consists in developing $u(\cdot, U)$ with the Itô formula and in writing it as the solution of a backward SDE. This permits to apply Gronwall arguments to prove that V matches $u(\cdot, U)$. This "decoupling strategy" seems to be relevant for equations of type $(\mathcal{E})$ that admit a strong solution. It has been applied to different frameworks: homogenization, see e.g. Buckdahn and Hu [6], and numerical approximation, see Delarue and Menozzi [10].

In the weak solvability framework, this so-called "decoupling method" still applies: most of the difficulty introduced by weakening the notion of solution consists in proving that uniqueness in law holds for (2.3). Thanks to the large literature devoted to the weak solvability of SDEs, this task is easily performed. Note at the opposite that the strategy proposed in Delarue [7], which consists in estimating the differences $X - U$ and $(Y - V, Z - W)$, completely fails for weak solutions: processes X and U are now defined on different probability spaces and there is no way to investigate the distance between them.

A priori *Estimates.* To be precise, note that the most difficult point in our setting consists in applying Gronwall's lemma to complete the "decoupling strategy". In short, this is possible if the partial derivatives of order one and two of the previous solution u are efficiently controlled. In Ma, Protter and Yong [30], in Buckdahn and Hu [6], or in Delarue and Menozzi [10], these derivarives are uniformly bounded on the whole space. In Guatteri and Lunardi [16], $|\nabla^2_{x,x} u(t, \cdot)|$ is locally bounded by $C(T - t)^{-(1-\gamma)}$, $\gamma > 0$. In our case, the story is rather different: the solution u does not belong to $\mathcal{C}^{1,2}([0, T[\times \mathbb{R}^d, \mathbb{R})$ but to $W^{1,2,p}_{\text{loc}}([0, T[\times \mathbb{R}^d, \mathbb{R})$ and there is no hope to obtain a pointwise control of the second order derivatives of u. The strategy then consists in a tricky application of the Krylov inequalities (see Krylov [21], Chapter II, Sections 2 and 3).

3 *A priori* Estimates in the Smooth Case

In this section, we assume that the coefficients are smooth, *i.e.* that they are infinitely differentiable with respect to the variables t, x, y and z, and bounded, with bounded derivatives of any order. As already explained in Subsection 2.2, it is then well-known that the quasi-linear PDE $(\mathcal{E})$ admits a unique solution $u \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d, \mathbb{R})$. Moreover, for an arbitrarily chosen standard set-up $(\Omega, \{\mathcal{F}\}, \mathbb{P}, B)$ (e.g. the canonical Wiener space), the FBSDE (E) admits a unique strong solution. For every initial condition $(t, x) \in [0, T] \times \mathbb{R}^d$, we denote this solution by $(X^{t,x}, Y^{t,x}, Z^{t,x})$. The triple $(X^{t,x}, Y^{t,x}, Z^{t,x})$ and the solution u are then connected by relationships (2.1) and (2.2).

We then present several *a priori* bounds of the solution u and of its derivatives in terms of known parameters appearing in Assumption **(A)**. These controls permit both to introduce a regularization procedure to prove the existence of a solution to $(\mathcal{E})$ under Assumption **(A)** and to apply the "decoupling strategy" to prove the weak unique solvability of (E).

3.1 *Supremum* Norm of u

We first give a probabilistic proof of the following estimate of the *supremum* norm of u:

Theorem 3.1 *There exists a constant $\Gamma_{3.1}$, depending only on d, Λ, and T, such that:*

$$\forall (t, x) \in [0, T] \times \mathbb{R}^d, \ |u(t, x)| \leq \Gamma_{3.1}.$$

Proof. The strategy is clear. We aim to show that there exists a constant $\Gamma_{3.1}$, depending only on Λ and T, such that for any initial condition $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$\mathbb{P}\text{−a.s.}, \ \|Y^{t,x}\|_\infty \le \Gamma_{3.1} \text{ and } \mathbb{E}\int_t^T |Z_s^{t,x}|^2 ds \le \Gamma_{3.1}. \tag{3.1}$$

Connection (2.2) and inequality (3.1) then permit to complete the proof.

Estimate (3.1) follows from Proposition 2.1 and Corollary 2.2, given in Kobylanski [20], with:

$$\begin{aligned}
a_0(t,v,z) &= \frac{f(t,x,v,z)}{1+|v|+|z|^2}\mathrm{sgn}(v),\\
F_0(t,v,z) &= \frac{f(t,x,v,z)}{1+|v|+|z|^2}(1+|z|^2),\\
a(t) &= \Lambda, \ b(t) = \Lambda, \ C = \Lambda.
\end{aligned}$$

This completes the proof. □

3.2 Hölder Estimate of u

According to Theorem 1.3 given in Delarue [8], we claim:

Theorem 3.2 *There exist a constant $\alpha_1 > 0$, depending only on d, λ and Λ, and a constant $\Gamma_{3.2}$, depending only on α_0, d, H, λ, Λ and T, such that :*

$$\forall (t,x), \ \forall (s,y) \in [0,T] \times \mathbb{R}^d, \ |u(t,x) - u(s,y)| \le \Gamma_{3.2}\big(|x-y|^{\alpha_2} + |t-s|^{\alpha_2/2}\big), \ \alpha_2 = \alpha_0 \wedge \alpha_1.$$

Say a word about the proof of Theorem 3.2. Recall in particular that the main argument derives from the Krylov and Safonov theory (see Krylov and Safonov [22] and [23] or Bass [4]). In short, this approach permits to establish the *a priori* Hölder continuity of the solutions to a linear parabolic PDE with a non-degenerate, but discontinuous, diffusion matrix.

The reader may object that Theorem 1.3 in [8] just holds for a coefficient f with linear growth in z. This is right: in [8], the backward process Y is multi-dimensional and, for this reason, the backward driver f cannot be quadratic in z. Nevertheless, the crucial starting point in the proof of Theorem 1.3 in [8] is the inequality (1.10) which permits to compare the backward process Y, or at least a variant of it, to the solution of the one-dimensional BSDE given in (1.12) in [8]. This is the reason why the first author focuses in [8] on $\mu(Y_s)_i + |Y_s|^2$ and not on $(Y_s)_i$ itself. In our current setting, this procedure is useless since the process Y can be directly compared to the solution of a quadratic BSDE of the same form as (1.12) in [8]. The issue of this comparison method clearly appears in (1.23) in [8]. In the end, the strategy used to prove Theorem 1.3 in [8] also applies under Assumption **(A)**.

3.3 *Supremum* Norm of the Gradient

The following estimate permits to bound the coefficients b and f in (E) and $(\mathcal{E})$:

Theorem 3.3 *There exist two constants $\alpha_3 > 0$ and $\Gamma_{3.3}$, depending only on $\alpha_0, d, H, \lambda, \Lambda$ and T, such that:*

$$\forall (t,x) \in [0,T[\times \mathbb{R}^d, \ |\nabla_x u(t,x)| \le \Gamma_{3.3}(T-t)^{-(1-\alpha_3)/2}.$$

The proof is given in Subsection 7.1.

3.4 Hölder Estimate of the Gradient

The following Hölder estimate of the gradient of u is proved in Subsection 7.2:

Theorem 3.4 *There exist two constant $\alpha_4 > 0$ and $\Gamma_{3.4}$, depending only on $\alpha_0, d, H, \lambda, \Lambda$ and T, such that :*

$$\begin{aligned}&\forall (t,x),(s,y) \in [0,T[\times\mathbb{R}^d,\ 0 \le t \le s < T,\\ &\qquad\qquad |\nabla_x u(t,x) - \nabla_x u(s,y)| \le \Gamma_{3.4}(T-s)^{-(1-\alpha_4)/2}\big(|x-y|^{\alpha_4} + |t-s|^{\alpha_4/2}\big).\end{aligned}$$

3.5 Calderón-Zygmund Estimates

Thanks to the Calderón-Zygmund inequalities, we prove in Section 6 the following L^p_{loc}-controls of $\partial_t u$ and $\nabla^2_{x,x} u$:

Theorem 3.5 *There exists a constant $\alpha_5 \in]0,1]$, depending only on $\alpha_0, d, H, \lambda, \Lambda$ and T, such that:*

$$\begin{aligned}&\forall p \ge 1,\ R \ge 1,\ \delta \in]0,T],\ z \in \mathbb{R}^d,\\ &\int_{T-\delta}^{T}\int_{B(z,R)} \big[(T-s)^{1-\alpha_5}\big(|\partial_t u(s,y)| + |\nabla^2_{x,x} u(s,y)|\big)\big]^p ds\, dy \le C_{3.5}(p)\delta R^d,\end{aligned}$$

where $C_{3.5}(p)$ depends only on $\alpha_0, d, H, \lambda, \Lambda, p$ and T.

4 Solvability of (E) and $(\mathcal{E})$

We now turn to the proof of Theorem 2.1.

4.1 Solvability of $(\mathcal{E})$

Thanks to Assumption **(A)**, we can consider a sequence $(b_n, f_n, \sigma_n, G_n)_{n\ge 1}$ of smooth coefficients, satisfying Assumption **(A)** with respect to α_0, H, CK, λ and $C\Lambda$, for a suitable universal constant $C > 0$, and converging towards (b, f, σ, G) in the following sense (as $n \to +\infty$):

$$\begin{cases}\text{For a.e. } (t,x) \in [0,T]\times\mathbb{R}^d,\ \forall (y,z) \in \mathbb{R}\times\mathbb{R}^d,\ (a_n \equiv \sigma_n\sigma_n^*, b_n, f_n)(t,x,y,z) \to (a,b,f)(t,x,y,z),\\ G_n \to G \text{ uniformly on compact subsets of } \mathbb{R}^d.\end{cases}$$

The reader can find a possible construction of these functions, up to the discontinuity of a in t, in Delarue [7] (the problem is in fact easier in our case since we just regularize a and not σ). Hence, for every $n \ge 1$, we can associate to the coefficients $(b_n, f_n, \sigma_n, G_n)$ a smooth solution u_n. Thanks to Theorems 3.1, 3.2, 3.3, 3.4 and 3.5, we can extract a subsequence, still indexed by n, such that u_n (resp. $\nabla_x u_n$) converges in *supremum* norm on every compact subset of $[0,T]\times\mathbb{R}^d$ (resp. $[0,T[\times\mathbb{R}^d$) and $\nabla^2_{x,x}u_n$ and $\partial_t u_n$ converge weakly, for every $\delta \in]0,1[$ and $p > 1$, in $L^p([0,T(1-\delta)]\times B(0,1/\delta),\mathbb{R}^{d\times d})$ and $L^p([0,T(1-\delta)]\times B(0,1/\delta),\mathbb{R})$. We denote by u the limit function. It is then clear that u satisfies almost everywhere the PDE $(\mathcal{E})$. Moreover, inequalities given in Theorems 3.1, 3.2, 3.3, 3.4 and 3.5 still hold. In particular, u is bounded and continuous on $[0,T]\times\mathbb{R}^d$.

4.2 Existence of a Weak Solution to (E)

This subsection is devoted to the weak solvability of (E): the initial condition is chosen to be of the form $(0,x)$, $x \in \mathbb{R}^d$ (of course, the proof applies to an initial condition of the form (t,x)). We adapt to this end the famous theory of Stroock and Varadhan [37].

Consider first on $[0,T]$ the martingale problem associated to $(0, a(\cdot,\cdot,u(\cdot,\cdot)))$. The diffusion coefficient $a(\cdot,\cdot,u(\cdot,\cdot))$ is, thanks to Assumptions **(A.1)** and **(A.2)**, and to Theorems 3.1 and 3.2, bounded, non-degenerate and continuous in x, uniformly in t. Referring to Theorem 7.2.1, Chapter VII in Stroock and Varadhan [37], this martingale problem is well-posed.

Focus now on the martingale problem associated to $(b(\cdot,\cdot,u(\cdot,\cdot),\nabla_x u(\cdot,\cdot)), a(\cdot,\cdot,u(\cdot,\cdot)))$ on $[0,T]$. The drift b is not bounded and thus does not fulfill the assumptions of Theorem 7.2.1, Chapter VII in Stroock and Varadhan [37]. However, according to Assumption **(A.1)** and to Theorems 3.1 and 3.3 (boundedness of u and local boundedness of $\nabla_x u$), the function $b(t,\cdot,u(t,\cdot),\nabla_x u(t,\cdot))$, for $t \in [0,T[$, is bounded by $\Lambda(1+\Gamma_{3.1}+\Gamma_{3.3}(T-t)^{-(1-\alpha_3)/2})$. This permits to apply the Girsanov transform to deduce the well-posedness of the current martingale problem from the well-posedness of the problem associated to $(0, a(\cdot,\cdot,u(\cdot,\cdot)))$ as done in Theorem 6.4.3, Chapter VI in Stroock and Varadhan [37] (see also Theorem 27.1, Chapter V in Rogers and Williams [36]). In our current setting, the Girsanov transform derives from the Novikov property (*cf.* Paragraph D, Section 3, Chapter III in Karatzas and Shreve [19], see also Subsection 37, Chapter IV in Rogers and Williams [36]). In particular, the SDE associated on $[0,T]$ to $(b(\cdot,\cdot,u(\cdot,\cdot),\nabla_x u(\cdot,\cdot)), \sigma(\cdot,\cdot,u(\cdot,\cdot)))$ and to the initial condition $(0,x)$ is uniquely solvable in the weak sense. Thus, there exists a standard set-up $(\Omega, \{\mathcal{F}\}, \mathbb{P}, B)$ and a continuous and $\{\mathcal{F}_t\}_{0\le t\le T}$-adapted process X with values in $\mathbb{R}^d$ such that:

$$\forall t \in [0,T],\ X_t = x + \int_0^t b\big(s, X_s, u(s,X_s), \nabla_x u(s,X_s)\big)ds + \int_0^t \sigma\big(s, X_s, u(s,X_s)\big)dB_s. \tag{4.1}$$

Thanks to **(A.1)** (growth of the coefficients), Theorems 3.1 (boundedness of u) and 3.3 (local boundedness of $\nabla_x u$), we claim:

$$\mathbb{E}\big[\sup_{t\in[0,T]} |X_t|^2\big] < +\infty. \tag{4.2}$$

Turn now to the backward equation and apply to this end the so-called Itô-Krylov formula (see Krylov [21], Chapter II, Section 10, Theorem 1) to the process (Y,Z) defined by:

$$\forall t \in [0,T], Y_t \equiv u(t,X_t),\ \forall t \in [0,T[,\ Z_t \equiv \nabla_x u(t,X_t).$$

For every $R > 0$, set $\rho(R) \equiv \inf\{t \ge 0,\ |X_t| \ge R\} \wedge T(1-1/R)$. The drift b in (4.1) is then bounded up to time $\rho(R)$ (see Theorem 3.3). The Itô-Krylov formula yields:

$$\forall 0 \le t \le \rho(R),\ Y_t = Y_{\rho(R)} + \int_t^{\rho(R)} f(s,X_s,Y_s,Z_s)ds - \int_t^{\rho(R)} \langle Z_s, \sigma(s,X_s,Y_s)dB_s\rangle.$$

Let R tend to $+\infty$: thanks to (4.2), $\rho(R) \to T$. Thanks to the continuity of the function u and of the process X in T, $Y_{\rho(R)}$ converges $\mathbb{P}$-a.s. towards $G(X_T)$. Due to the boundedness of u (Theorem 3.1) and to the control of $\nabla_x u$ (see Theorem 3.3), the driver of the backward equation is, $\mathbb{P}$-a.s., integrable over $[0,T]$. Finally, according again to Theorem 3.3, the martingale part is square-integrable under $\mathbb{P}$. We deduce that :

$$\forall 0 \le t \le T,\ Y_t = G(X_T) + \int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T \langle Z_s, \sigma(s,X_s,Y_s)dB_s\rangle,$$

with,

$$\mathbb{E}\int_0^T |Z_s|^2 ds < +\infty.$$

Moreover, thanks again to Theorem 3.1, there exists a constant $C \ge 0$ such that $\mathbb{P}-$a.s.:

$$\sup_{t\in[0,T]} |Y_t| \le C.$$

Hence the triple of $\{\mathcal{F}_t\}_{0\le t\le T}$ progressively-measurable processes (X,Y,Z) together with the set-up $(\Omega,\{\mathcal{F}\},\mathbb{P},B)$ is a weak solution of (E) with initial condition $(0,x)$. This proves the weak solvability of (E). $\square$

4.3 Uniqueness in law

We now focus on the uniqueness in law of the solution (the initial condition $(0,x)$, $x\in\mathbb{R}^d$, being fixed).

4.3.1 Strategy

Recall that (X,Y,Z) denotes the solution built in Subsection 4.2 (with initial condition $(0,x)$, $x\in\mathbb{R}^d$) and consider another solution to the FBSDE (E) with the same initial condition: (U,V,W) with standard set-up $(\tilde{\Omega},\{\tilde{\mathcal{F}}\},\tilde{\mathbb{P}},\tilde{B})$. Set also:

$$\forall t\in[0,T],\ \bar{V}_t\equiv u(t,U_t),\ \forall t\in[0,T[,\ \bar{W}_t\equiv\nabla_x u(t,U_t).$$

The strategy aims to identify $(\bar{V},\bar{W})$ with (V,W): this permits to identify the forward component of (E) with the SDE satisfied by X (see (4.1)), and thus to derive Theorem 2.1 from the weak uniqueness property of (4.1).

The proof is divided in several steps. We first apply the Itô-Krylov formula to the process $\bar{V}$ to write it as the solution of a backward equation. Using a suitable quadratic functional, we then investigate the difference between (V,W) and $(\bar{V},\bar{W})$. Thanks to the Krylov estimates and to Theorem 3.5, we prove that the difference $V-\bar{V}$ satisfies a non-standard discrete Gronwall inequality. We finally derive that (V,W) matches $(\bar{V},\bar{W})$.

4.3.2 Girsanov Change of Measure and Itô-Krylov Formula

We first aim to apply the Itô formula to the quantity $u(\cdot,U)$. Unfortunately, the function u does not belong to $\mathcal{C}^{1,2}([0,T[\times\mathbb{R}^d,\mathbb{R})$ since the partial derivatives of u of order one in t and of order two in x are just defined in the Sobolev sense (*cf.* Subsection 4.1). Hence, the classical Itô formula does not apply. To overcome the lack of regularity of u, we refer again to the Itô-Krylov formula (*cf.* Krylov [21], Chapter II, Section 10, Theorem 1). Roughly speaking, if the drift b of the Itô process U is bounded, the process $\bar{V}$ still develops as a semi-martingale.

The point is that the *supremum* norm of b is not finite in our frame since the process W is not bounded. We thus change the underlying probability measure to get rid of the drift b in the writing of U.

Fix to this end a real $A>0$ and define $\zeta(A)\equiv\inf\Big\{t\ge 0,\ \int_0^t|W_s|^2ds>A\Big\}\wedge T$.

According to the well-known Novikov condition (*cf.* Paragraph D, Section 3, Chapter III in Karatzas and Shreve [19]), the process $\bar{B}$ given by:

$$\forall t\in[0,T],\ \bar{B}_t\equiv\tilde{B}_t+\int_0^{t\wedge\zeta(A)}\sigma^{-1}(s,U_s,V_s)b(s,U_s,V_s,W_s)ds, \tag{4.3}$$

is an $\{\tilde{\mathcal{F}}_t\}_{0\le t\le T}$-Brownian motion under the probability $\bar{\mathbb{P}}$ given by:

$$\frac{d\bar{\mathbb{P}}}{d\tilde{\mathbb{P}}}\equiv\exp\Big(-\int_0^{\zeta(A)}\langle\sigma^{-1}(s,U_s,V_s)b(s,U_s,V_s,W_s),d\tilde{B}_s\rangle\Big)$$
$$\times\exp\Big(-\frac{1}{2}\int_0^{\zeta(A)}\big|\sigma^{-1}(s,U_s,V_s)b(s,U_s,V_s,W_s)\big|^2ds\Big).$$

Mention carefully that the measures $\tilde{\mathbb{P}}$ and $\bar{\mathbb{P}}$ are equivalent. Due to Theorem 3.1 (boundedness of u) and to Definition 2.2 ($V \in S^{\infty}_{0,T}(\tilde{\Omega}, \{\tilde{\mathcal{F}}\}, \tilde{\mathbb{P}}, \mathbb{R})$), processes V and $\bar{V}$ are almost-surely bounded under the new probability $\bar{\mathbb{P}}$.

Define then the following process:

$$\forall t \in [0,T], \ \bar{U}_t \equiv x + \int_0^t \sigma(s, U_s, V_s) d\bar{B}_s. \tag{4.4}$$

According to (4.3) and (4.4), note that $\bar{U}$ and U match on $[0, \zeta(A)]$:

$$\forall t \in [0, \zeta(A)], \ \bar{U}_t = U_t. \tag{4.5}$$

As done in the former subsection, consider also, for a given real $R > 0$, the stopping time $\bar{\rho}(R) \equiv \inf\{t \geq 0, \ |\bar{U}_t| \geq R\} \wedge T(1 - 1/R)$. In short, $\bar{\rho}(R)$ permits to localize the values of the process $\bar{U}$ and thus to apply the Itô-Krylov formula to the process $u(\cdot, \bar{U})$.

Recall now from Subsection 4.1 that u belongs to $W^{1,2,d+1}_{\text{loc}}([0,T[\times \mathbb{R}^d, \mathbb{R})$. Hence, for a given stopping time $\tau \leq \zeta(A) \wedge \bar{\rho}(R)$, we claim from Krylov [21], Chapter II, Section 10, Theorem 1:

$$\begin{aligned}
&\bar{\mathbb{P}}-\text{a.s.}, \ \forall t \in [0,\tau], \ du(t, \bar{U}_t) \\
&\quad = \partial_t u(t, \bar{U}_t) dt + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(t, U_t, V_t) \partial^2_{x_i, x_j} u(t, \bar{U}_t) dt + \langle \nabla_x u(t, \bar{U}_t), \sigma(t, U_t, V_t) d\bar{B}_t \rangle.
\end{aligned}$$

Due to (4.5), note that we can replace $\bar{U}$ by U in the above equality. Hence, using that u is a solution of the equation $(\mathcal{E})$, we deduce for every $t \in [0, \tau]$:

$$\begin{aligned}
du(t, U_t) \ &= \frac{1}{2} \sum_{i,j=1}^d \left[a_{i,j}(t, U_t, V_t) - a_{i,j}(t, U_t, \bar{V}_t) \right] \partial^2_{x_i, x_j} u(t, U_t) dt \\
&\quad - \left[\langle b(t, U_t, \bar{V}_t, \bar{W}_t), \bar{W}_t \rangle + f(t, U_t, \bar{V}_t, \bar{W}_t) \right] dt \\
&\quad + \langle \bar{W}_t, \sigma(t, U_t, V_t) d\bar{B}_t \rangle.
\end{aligned} \tag{4.6}$$

Focus a while on the bounded variation terms in (4.6). The PDE $(\mathcal{E})$ and the Krylov inequalities (*cf.* Sections 2 and 3, Chapter II in Krylov [21]) ensure that the following terms make sense and are equal:

$$\begin{aligned}
\int_0^\tau \partial_t u(t, U_t) dt = - \int_0^\tau \Bigg[&\frac{1}{2} \sum_{i,j=1}^d \left[a_{i,j}(t, U_t, \bar{V}_t) \partial^2_{x_i, x_j} u(t, U_t) \right] \\
&+ \left[\langle b(t, U_t, \bar{V}_t, \bar{W}_t), \bar{W}_t \rangle + f(t, U_t, \bar{V}_t, \bar{W}_t) \right] \Bigg] dt.
\end{aligned}$$

In fact, due to Assumption **(A.1)** (growth of the coefficients), to Theorems 3.1 and 3.3 (boundedness of u and local boundedness of $\nabla_x u$), to the definition of τ and again to the Krylov inequalities, each dt-term in (4.6) is correctly defined.

Note now from (E) that dV writes:

$$\bar{\mathbb{P}}-\text{a.s.}, \ \forall t \in [0,\tau], \ dV_t = -\left[\langle b(t, U_t, V_t, W_t), W_t \rangle + f(t, U_t, V_t, W_t) \right] dt + \langle W_t, \sigma(t, U_t, V_t) d\bar{B}_t \rangle. \tag{4.7}$$

Therefore, from (4.6) and (4.7), we obtain for every $t \in [0, \tau]$:

$$\begin{aligned} d(V-\bar{V})_t &= -\frac{1}{2}\sum_{i,j=1}^{d}\big[a_{i,j}(t,U_t,V_t) - a_{i,j}(t,U_t,\bar{V}_t)\big]\partial^2_{x_i,x_j}u(t,U_t)dt \\ &\quad - \big[\langle b(t,U_t,V_t,W_t),W_t\rangle - \langle b(t,U_t,\bar{V}_t,\bar{W}_t),\bar{W}_t\rangle\big]dt \\ &\quad - \big[f(t,U_t,V_t,W_t) - f(t,U_t,\bar{V}_t,\bar{W}_t)\big]dt. \\ &\quad + \langle W_t - \bar{W}_t, \sigma(t,U_t,V_t)d\bar{B}_t\rangle \end{aligned} \tag{4.8}$$

Define for the sake of simplicity $\forall (t,x,y,z) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d$, $F(t,x,y,z) \equiv \langle b(t,x,y,z),z\rangle + f(t,x,y,z)$. Note that F satisfies a similar bound to (**A.5**):

$$\begin{aligned} &\big|F(t,x,y,z) - F(t,x,y',z')\big| \\ &\quad \le C_{4.0}\big(1+|y|+2|z|+|z'-z|\big)\big(|y-y'|+|z-z'|\big), \end{aligned} \tag{4.9}$$

with $C_{4.0} \equiv 2K + \Lambda$.

4.3.3 Quadratic Functional of $V - \bar{V}$

We now apply a variant of the method used in the proof of Proposition 2.1, in Kobylanski [20]. Set to this end $L = 2(\|V\|_\infty^2 + \|\bar{V}\|_\infty^2)$ (recall that V is $\tilde{\mathbb{P}}$ and $\bar{\mathbb{P}}$ almost-surely bounded and that u is bounded, *cf.* Theorem 3.1) and define the following function:

$$\Phi(z) = \exp(cz) - 1, \;\; z \in [0, L], \tag{4.10}$$

where c denotes a free nonnegative parameter whose value is chosen in the sequel. It is easy to show that $\Phi \in \mathcal{C}^2([0,L],\mathbb{R})$ and that for all $z \in [0,L]$:

$$\begin{cases} \text{(a) } \Phi(z) \ge 0 \text{ and } \Phi(z) = 0 \text{ iff } z = 0, \\ \text{(b) } c\exp(cL) \ge \Phi'(z) \ge c, \\ \text{(c) } cz \le z\Phi'(z) \le (cL+1)\Phi(z), \\ \text{(d) } c\Phi'(z) - \Phi''(z) = 0. \end{cases} \tag{4.11}$$

Apply Itô's formula to $\Phi(|V_\cdot - \bar{V}_\cdot|^2)$. Due to (4.8), for $0 \le t \le \tau$:

$$\begin{aligned} \Phi(|V_t - \bar{V}_t|^2) &= \Phi(|V_\tau - \bar{V}_\tau|^2) \\ &+ \int_t^\tau \Big[\Phi'(|V_s-\bar{V}_s|^2)[V_s-\bar{V}_s]\Big(\sum_{i,j=1}^{d}\big[a_{i,j}(s,U_s,V_s) - a_{i,j}(s,U_s,\bar{V}_s)\big]\partial^2_{x_i,x_j}u(s,U_s)\Big)\Big]ds \\ &+ 2\int_t^\tau \Big[\Phi'(|V_s-\bar{V}_s|^2)[V_s-\bar{V}_s]\big[F(s,U_s,V_s,W_s) - F(s,U_s,\bar{V}_s,\bar{W}_s)\big]\Big]ds \\ &- 2\int_t^\tau \Big[\Phi'(|V_s-\bar{V}_s|^2)[V_s-\bar{V}_s]\langle W_s-\bar{W}_s, \sigma(s,U_s,V_s)d\bar{B}_s\rangle\Big] \\ &- \int_t^\tau \Big[\Phi'(|V_s-\bar{V}_s|^2)\langle W_s-\bar{W}_s, a(s,U_s,V_s)(W_s-\bar{W}_s)\rangle\Big]ds \\ &- 2\int_t^\tau \Big[\Phi''(|V_s-\bar{V}_s|^2)[V_s-\bar{V}_s]^2\langle W_s-\bar{W}_s, a(s,U_s,V_s)(W_s-\bar{W}_s)\rangle\Big]ds. \end{aligned}$$

Taking into account Assumptions **(A.2)** and **(A.3)**, (4.9) (regularity of F), Theorems 3.1 (boundedness of u) and 3.3 (local boundedness of $\nabla_x u$), (4.11)-(b) ($\Phi' \ge 0$) and (4.11)-(d) ($c\Phi' - \Phi'' = 0$), there exists a constant $C > 0$, which may change from line to line but depends only on L and on known parameters appearing in **(A)**, such that for $0 \le t \le \tau$:

$$
\begin{aligned}
&\Phi(|V_t - \bar{V}_t|^2) + \lambda \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)\big(1 + 2c|V_s - \bar{V}_s|^2\big)|W_s - \bar{W}_s|^2\big] ds \\
&\le \Phi(|V_\tau - \bar{V}_\tau|^2) \\
&\quad + C \int_t^\tau \big[\big(1 + (T-s)^{-1/2+\alpha_3/2} + |\nabla^2_{x,x} u(s, U_s)|\big)\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s|^2\big] ds \\
&\quad + C \int_t^\tau \big[\big(1 + (T-s)^{-1/2+\alpha_3/2}\big)\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s||W_s - \bar{W}_s|\big] ds \\
&\quad + 2C_{4.0} \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s||W_s - \bar{W}_s|\big(|V_s - \bar{V}_s| + |W_s - \bar{W}_s|\big)\big] ds \\
&\quad - \int_t^\tau dM_s,
\end{aligned}
$$

with:

$$
\forall s \in [0, T],\ dM_s \equiv 2\mathbf{1}_{\{s \le \tau\}} \Phi'(|V_s - \bar{V}_s|^2)[V_s - \bar{V}_s]\langle W_s - \bar{W}_s, \sigma(s, U_s, V_s) d\bar{B}_s\rangle. \tag{4.12}
$$

From the classical Young inequality ($2ab \le ka^2 + k^{-1}b^2$, $k > 0$):

$$
\begin{aligned}
&\Phi(|V_t - \bar{V}_t|^2) + \lambda \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)\big(3/4 + 2c|V_s - \bar{V}_s|^2\big)|W_s - \bar{W}_s|^2\big] ds \\
&\le \Phi(|V_\tau - \bar{V}_\tau|^2) \\
&\quad + C \int_t^\tau \big[\big(1 + (T-s)^{-1+\alpha_3} + |\nabla^2_{x,x} u(s, U_s)|\big)\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s|^2\big] ds \\
&\quad + 2C_{4.0} \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s||W_s - \bar{W}_s|\big(|V_s - \bar{V}_s| + |W_s - \bar{W}_s|\big)\big] ds \\
&\quad - \int_t^\tau dM_s.
\end{aligned}
$$

Focus on the third term in the r.h.s of the above inequality. Use first the boundedness of V and $\bar{V}$ and then Young's inequality to deduce:

$$
\begin{aligned}
&\Phi(|V_t - \bar{V}_t|^2) + \lambda \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)\big(1/2 + 2c|V_s - \bar{V}_s|^2\big)|W_s - \bar{W}_s|^2\big] ds \\
&\le \Phi(|V_\tau - \bar{V}_\tau|^2) \\
&\quad + C \int_t^\tau \big[\big(1 + (T-s)^{-1+\alpha_3} + |\nabla^2_{x,x} u(s, U_s)|\big)\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s|^2\big] ds \\
&\quad + 2C_{4.0} \int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)|V_s - \bar{V}_s||W_s - \bar{W}_s|^2\big] ds \\
&\quad - \int_t^\tau dM_s.
\end{aligned}
$$

Apply again the classical Young inequality to the third term and deduce that:

$$
\begin{aligned}
&\Phi(|V_t-\bar V_t|^2)+\lambda\int_t^\tau \big[\Phi'(|V_s-\bar V_s|^2)\big(1/4+2c|V_s-\bar V_s|^2\big)|W_s-\bar W_s|^2\big]ds\\
&\le \Phi(|V_\tau-\bar V_\tau|^2)\\
&\quad + C\int_t^\tau \big[\big(1+(T-s)^{-1+\alpha_3}+|\nabla^2_{x,x}u(s,U_s)|\big)\Phi'(|V_s-\bar V_s|^2)|V_s-\bar V_s|^2\big]ds\\
&\quad + 4\lambda^{-1}C_{4.0}^2\int_t^\tau \big[\Phi'(|V_s-\bar V_s|^2)|V_s-\bar V_s|^2|W_s-\bar W_s|^2\big]ds\\
&\quad - \int_t^\tau dM_s.
\end{aligned}
$$

Choose now $c=2\lambda^{-2}C_{4.0}^2$ and deduce from (4.11)-(c) ($z\Phi'(z)\le (cL+1)\Phi(z)$) and from the boundedness of V and $\bar V$:

$$
\begin{aligned}
&\Phi(|V_t-\bar V_t|^2)+(\lambda/4)\int_t^\tau \big[\Phi'(|V_s-\bar V_s|^2)|W_s-\bar W_s|^2\big]ds\\
&\le \Phi(|V_\tau-\bar V_\tau|^2)\\
&\quad + C\int_t^\tau \big[\big(1+(T-s)^{-1+\alpha_3}+|\nabla^2_{x,x}u(s,U_s)|\big)\Phi(|V_s-\bar V_s|^2)\big]ds\\
&\quad - \int_t^\tau dM_s.
\end{aligned}
\tag{4.13}
$$

4.3.4 Krylov and Bernstein Inequalities

Focus on (4.13). The usual approach to identify (V,W) with $(\bar V,\bar W)$ (as developed in Pardoux and Peng [33] and in Ma, Protter and Yong [30]) consists in taking the expectation in (4.13) to apply a classical Gronwall argument. In short, this method holds when the second order derivatives of u are uniformly controlled on the whole set $[0,T]\times\mathbb{R}^d$ or, at least, locally bounded with an integrable singularity in the neighbourhood of the boundary T. As explained above, this point of view fails in our framework since Theorem 3.5 just provides an L^p estimate of $\nabla^2_{x,x}u$ (*cf.* Subsection 2.4)

To our own point of view, the most relevant argument to handle the r.h.s. in (4.13) derives again from the Krylov inequalities (see Sections 2 and 3, Chapter II in Krylov [21]). Roughly speaking, for every function $\ell\in L^{d+1}([0,T]\times\mathbb{R}^d)$:

$$
\bar{\mathbb{E}}\Big[\int_0^T |\ell(s,\bar U_s)|ds\Big]\le C\Big[\int_0^T\int_{\mathbb{R}^d}|\ell|^{d+1}(s,x)dsdx\Big]^{1/d+1}. \tag{4.14}
$$

Fix now $t\in[0,T[$, multiply both sides in (4.13) by $\mathbf{1}_{\{t\le\tau\}}$ and take the conditional expectation under $\bar{\mathbb{P}}$ with respect to $\tilde{\mathcal{F}}_t$. Due to the boundedness of V and $\bar V$, to Assumption **(A.1)** (growth of the coefficients), Theorem 3.3 (local boundedness of the gradient), to (4.11)-(b) ($\Phi'(z)\le c\exp(cL)$) and to the definition of $\zeta(A)$, the martingale part M (*cf.* (4.12)) is square-integrable, and, in particular, its conditional expectation vanishes:

$$
\begin{aligned}
&\mathbf{1}_{\{t\le\tau\}}\Phi(|V_t-\bar V_t|^2)+\mathbf{1}_{\{t\le\tau\}}(\lambda/4)\bar{\mathbb{E}}\Big[\int_t^\tau\big[\Phi'(|V_s-\bar V_s|^2)|W_s-\bar W_s|^2\big]ds\,\big|\tilde{\mathcal{F}}_t\Big]\\
&\le \mathbf{1}_{\{t\le\tau\}}\bar{\mathbb{E}}\Big[\Phi(|V_\tau-\bar V_\tau|^2)\,\big|\tilde{\mathcal{F}}_t\Big]\\
&\quad + C\mathbf{1}_{\{t\le\tau\}}\bar{\mathbb{E}}\Big[\int_t^\tau\big[\big(1+(T-s)^{-1+\alpha_3}+|\nabla^2_{x,x}u(s,\bar U_s)|\big)\Phi(|V_s-\bar V_s|^2)\big]ds\,\big|\tilde{\mathcal{F}}_t\Big].
\end{aligned}
\tag{4.15}
$$

Choose $\tau = \tau(t,r) \wedge \bar{\rho}(R) \wedge \zeta(A)$, with $r \geq 1$ and $\tau(t,r) \equiv \inf\{s \geq t,\ |\bar{U}_s - \bar{U}_t| \geq r\} \wedge T$, and admit for the moment the following Lemma:

Lemma 4.1 *There exists a constant $\gamma \in]0,1]$, depending only on known parameters appearing in Assumption* **(A)**, *such that for all $p \geq 1$ and $r \geq 1$:*

$$\bar{\mathbb{P}} - \text{a.s.},\ \bar{\mathbb{E}}\bigg[\int_t^{\tau(t,r)} \big[(T-s)^{(1-\gamma)p} |\nabla^2_{x,x} u(s,\bar{U}_s)|^p\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg] \leq C_{4.1}(p)(T-t)^{1/(d+1)} r^{d/d+1},$$

where $C_{4.1}(p)$ depends only on p and on known parameters in **(A)**.

Note that we can assume without loss of generality that $\gamma \leq \alpha_3$ (*cf.* α_3 in (4.15)). Write now $|\nabla^2_{x,x} u(s,\bar{U}_s)| = (T-s)^{-(1-\gamma)}(T-s)^{1-\gamma}|\nabla^2_{x,x} u(s,\bar{U}_s)|$ in (4.15) and apply the general Young inequality ($ab \leq a^q/q + b^p/p,\ 1/q + 1/p = 1$) with $q = (1-\gamma/2)/(1-\gamma)$ and $p = 2(1-\gamma/2)/\gamma$:

$$\begin{aligned}
&\mathbf{1}_{\{t \leq \tau\}} \Phi(|V_t - \bar{V}_t|^2) + \mathbf{1}_{\{t \leq \tau\}} (\lambda/4) \bar{\mathbb{E}}\bigg[\int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)|W_s - \bar{W}_s|^2\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg] \\
&\leq \mathbf{1}_{\{t \leq \tau\}} \bar{\mathbb{E}}\Big[\Phi(|V_\tau - \bar{V}_\tau|^2)|\tilde{\mathcal{F}}_t\Big] \\
&\quad + C \mathbf{1}_{\{t \leq \tau\}} \bar{\mathbb{E}}\bigg[\int_t^\tau \big[\big(1 + (T-s)^{-(1-\gamma)} + (T-s)^{-(1-\gamma)(1-\gamma/2)/(1-\gamma)} \\
&\qquad + (T-s)^{2(1-\gamma)(1-\gamma/2)/\gamma} |\nabla^2_{x,x} u(s,\bar{U}_s)|^{2(1-\gamma/2)/\gamma}\big) \Phi(|V_s - \bar{V}_s|^2)\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg]. \\
&\leq \mathbf{1}_{\{t \leq \tau\}} \bar{\mathbb{E}}\Big[\Phi(|V_\tau - \bar{V}_\tau|^2)\, \big|\tilde{\mathcal{F}}_t\Big] \\
&\quad + C \mathbf{1}_{\{t \leq \tau\}} \bar{\mathbb{E}}\bigg[\int_t^{\tau(t,r)} \big[\big(1 + (T-s)^{-(1-\gamma/2)} \\
&\qquad + (T-s)^{2(1-\gamma)(1-\gamma/2)/\gamma} |\nabla^2_{x,x} u(s,\bar{U}_s)|^{2(1-\gamma/2)/\gamma}\big) \Phi(|V_s - \bar{V}_s|^2)\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg].
\end{aligned}$$

Apply Lemma 4.1 with $p = 2(1-\gamma/2)/\gamma$:

$$\begin{aligned}
&\mathbf{1}_{\{t \leq \tau\}} \Phi(|V_t - \bar{V}_t|^2) + \mathbf{1}_{\{t \leq \tau\}} (\lambda/4) \bar{\mathbb{E}}\bigg[\int_t^\tau \big[\Phi'(|V_s - \bar{V}_s|^2)|W_s - \bar{W}_s|^2\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg] \\
&\leq \mathbf{1}_{\{t \leq \tau\}} \bar{\mathbb{E}}\big[\Phi(|V_\tau - \bar{V}_\tau|^2)\, \big|\tilde{\mathcal{F}}_t\big] \\
&\quad + C\big[(T-t)^{\gamma/2} + (T-t)^{1/(d+1)} r^{d/(d+1)}\big] \sup_{t \leq s \leq T} \operatorname*{essup}_{\omega \in \Omega} \big[\Phi(|V_s - \bar{V}_s|^2)\big].
\end{aligned} \tag{4.16}$$

Recall that the values of the above essential *suprema* are the same under $\tilde{\mathbb{P}}$ and under $\bar{\mathbb{P}}$ since these measures are equivalent.

The strategy now consists in letting $R \to +\infty$ in (4.16). This is rather easy since $\bar{\mathbb{P}}$ does not depend on R. Note indeed that $\sup_{0 \leq s \leq T} |\bar{U}_s|$ belongs to $L^2(\Omega, \bar{\mathbb{P}})$ and deduce in particular that, $\bar{\mathbb{P}}$ almost-surely, $\bar{\rho}(R) \to T$ as $R \to +\infty$ (*cf.* Subsection 4.3.2 for the definitions of $\bar{U}$ and $\bar{\rho}(R)$). Hence, $\tau \to \tau_\infty \equiv \tau(t,r) \wedge \zeta(A)$ as $R \to +\infty$ (*cf.* the lines preceding Lemma 4.1 for the definitions of τ and $\tau(t,r)$).

Since V and $\bar{V}$ are bounded and continuous and since Φ is smooth, (4.16) yields:

$$\begin{aligned}
&\mathbf{1}_{\{t \leq \tau_\infty\}} \Phi(|V_t - \bar{V}_t|^2) + \mathbf{1}_{\{t \leq \tau_\infty\}} (\lambda/4) \bar{\mathbb{E}}\bigg[\int_t^{\tau_\infty} \big[\Phi'(|V_s - \bar{V}_s|^2)|W_s - \bar{W}_s|^2\big] ds\, \big|\tilde{\mathcal{F}}_t\bigg] \\
&\leq \mathbf{1}_{\{t \leq \tau_\infty\}} \bar{\mathbb{E}}\big[\Phi(|V_{\tau_\infty} - \bar{V}_{\tau_\infty}|^2)\, \big|\tilde{\mathcal{F}}_t\big] \\
&\quad + C\big[(T-t)^{\gamma/2} + (T-t)^{1/(d+1)} r^{d/(d+1)}\big] \sup_{t \leq s \leq T} \operatorname*{essup}_{\omega \in \Omega} \big[\Phi(|V_s - \bar{V}_s|^2)\big].
\end{aligned} \tag{4.17}$$

Since $\bar{\mathbb{P}}$ does depend on A, the same strategy fails to investigate the asymptotic form of (4.17) as $A \to +\infty$. We thus need to focus more precisely on (4.17) and in particular on the first term in the r.h.s:

$$
\begin{aligned}
&\bar{\mathbb{E}}\big[\Phi(|V_{\tau_\infty} - \bar{V}_{\tau_\infty}|^2)\,\big|\tilde{\mathcal{F}}_t\big] \\
&\quad \le C\bar{\mathbb{P}}\big\{\tau_\infty < T\,\big|\tilde{\mathcal{F}}_t\big\} + \bar{\mathbb{E}}\big[\Phi(|V_T - \bar{V}_T|^2)\,\big|\tilde{\mathcal{F}}_t\big] \\
&\quad \le C\bar{\mathbb{P}}\big\{\tau_\infty = \zeta(A), \zeta(A) < T\,\big|\tilde{\mathcal{F}}_t\big\} + C\bar{\mathbb{P}}\big\{\tau(t,r) < T\,\big|\tilde{\mathcal{F}}_t\big\} + \bar{\mathbb{E}}\big[\Phi(|V_T - \bar{V}_T|^2)\,\big|\tilde{\mathcal{F}}_t\big] \\
&\quad \equiv T(1) + T(2) + T(3).
\end{aligned}
\tag{4.18}
$$

The reader may object that $V_T - \bar{V}_T$ reduces to zero. In fact, we aim to keep the form written in (4.18) to derive later a crucial induction principle. Note now from the definition of $\zeta(A)$ (*cf.* Subsection 4.3.2) and from Theorem 3.3 (estimate of $\nabla_x u$) and (4.11)-(b) ($\Phi' \ge c$) that:

$$
\begin{aligned}
T(1) &\le C\bar{\mathbb{P}}\Big\{\int_0^{\tau_\infty} |W_s|^2 ds \ge A\,\big|\tilde{\mathcal{F}}_t\Big\} \\
&\le C\bar{\mathbb{P}}\Big\{t \le \tau_\infty,\ \int_0^{\tau_\infty} |W_s|^2 ds \ge A\,\big|\tilde{\mathcal{F}}_t\Big\} + C\mathbf{1}_{\{\tau_\infty < t\}} \\
&\le C\bar{\mathbb{P}}\Big\{t \le \tau_\infty,\ \int_t^{\tau_\infty} |W_s|^2 ds \ge A/2\,\big|\tilde{\mathcal{F}}_t\Big\} + C\mathbf{1}_{\{\zeta(A/2) \le t\}} + C\mathbf{1}_{\{\tau_\infty < t\}} \\
&\le C\bar{\mathbb{P}}\Big\{t \le \tau_\infty,\ \int_t^{\tau_\infty} |W_s - \bar{W}_s|^2 ds \ge A/8\,\big|\tilde{\mathcal{F}}_t\Big\} \\
&\quad + C\bar{\mathbb{P}}\Big\{t \le \tau_\infty,\ \int_t^{\tau_\infty} |\bar{W}_s|^2 ds \ge A/8\,\big|\tilde{\mathcal{F}}_t\Big\} + C\mathbf{1}_{\{\zeta(A/2) \le t\}} + C\mathbf{1}_{\{\tau_\infty < t\}} \\
&\le \mathbf{1}_{\{t \le \tau_\infty\}} C A^{-1} \bar{\mathbb{E}}\Big[\int_t^{\tau_\infty} \big[\Phi'(|V_s - \bar{V}_s|^2)|W_s - \bar{W}_s|^2\big] ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\quad + C\mathbf{1}_{\{A \le C\}} + C\mathbf{1}_{\{\zeta(A/2) \le t\}} + C\mathbf{1}_{\{\tau_\infty < t\}}.
\end{aligned}
\tag{4.19}
$$

Note now from the definition of $\tau(t,r)$ (*cf.* the lines preceding Lemma 4.1) and from the well-known Bernstein inequality, see e.g. Theorem IV.37.8 in Rogers and Williams [36] (note that the result holds true with such a conditional probability):

$$
T(2) \le \bar{\mathbb{P}}\Big\{\sup_{t \le s \le T}\Big|\int_t^s \sigma(s, U_s, V_s) d\bar{B}_s\Big| \ge r\,\big|\tilde{\mathcal{F}}_t\Big\} \le C\exp(-C^{-1} r^2 (T-t)^{-1}).
\tag{4.20}
$$

Derive from (4.18), (4.19) and (4.20):

$$
\begin{aligned}
\bar{\mathbb{E}}\big[\Phi(|V_{\tau_\infty} - \bar{V}_{\tau_\infty}|^2)\,\big|\tilde{\mathcal{F}}_t\big] &\le C\mathbf{1}_{\{\tau_\infty < t\}} + C\mathbf{1}_{\{A \le C\}} + C\mathbf{1}_{\{\zeta(A/2) \le t\}} \\
&\quad + \mathbf{1}_{\{t \le \tau_\infty\}} C A^{-1} \bar{\mathbb{E}}\Big[\int_t^{\tau_\infty} \big[\Phi'(|V_s - \bar{V}_s|^2)|W_s - \bar{W}_s|^2\big] ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\quad + C\exp(-C^{-1} r^2 (T-t)^{-1}) + \bar{\mathbb{E}}\big[\Phi(|V_T - \bar{V}_T|^2)\,\big|\tilde{\mathcal{F}}_t\big].
\end{aligned}
\tag{4.21}
$$

Thus, for A greater than a universal constant C', derive from (4.17) and (4.21):

$$
\begin{aligned}
\mathbf{1}_{\{t \le \tau_\infty\}} \Phi(|V_t - \bar{V}_t|^2) &\le \operatorname*{essup}_{\omega \in \Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big] + C\exp(-C^{-1} r^2 (T-t)^{-1}) + C\mathbf{1}_{\{\zeta(A/2) \le t\}} \\
&\quad + C\big[(T-t)^{\gamma/2} + (T-t)^{1/(d+1)} r^{d/(d+1)}\big] \sup_{t \le s \le T} \operatorname*{essup}_{\omega \in \Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big].
\end{aligned}
\tag{4.22}
$$

Mention carefully that (4.22) holds $\bar{\mathbb{P}}$ almost-surely and thus holds also $\tilde{\mathbb{P}}$ almost-surely.

4.3.5 Discrete Gronwall's Lemma

Let $A \to +\infty$. Then, $\tilde{\mathbb{P}}$ almost-surely, $\zeta(A), \zeta(A/2) \to T$ and $\tau_\infty \to \tau(t,r)$ (note that W belongs to $L^2(\Omega \times [0,T], \tilde{\mathbb{P}} \otimes ds)$). Since (4.22) holds $\tilde{\mathbb{P}}$ almost-surely, derive that:

$$\begin{aligned}\Phi(|V_t - \bar{V}_t|^2) \le \operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big] + C\exp(-C^{-1}r^2(T-t)^{-1})\\ + C\big[(T-t)^{\gamma/2} + (T-t)^{1/(d+1)}r^{d/(d+1)}\big] \sup_{t\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big].\end{aligned}$$

Note that the same inequality holds for every $t' \in [t,T]$. Thus:

$$\begin{aligned}\sup_{t\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big] \le \operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big] + C\exp(-C^{-1}r^2(T-t)^{-1})\\ + C\big[(T-t)^{\gamma/2} + (T-t)^{1/(d+1)}r^{d/(d+1)}\big] \sup_{t\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big].\end{aligned}$$

Choose now $r = (T-t)m$, $m \ge (T-t)^{-1}$ (to ensure $r \ge 1$ as required in Lemma 4.1):

$$\begin{aligned}\sup_{t\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big] \le \operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big] + C\exp(-C^{-1}m^2(T-t))\\ + C\big[(T-t)^{\gamma/2} + (T-t)m^{d/(d+1)}\big] \sup_{t\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big].\end{aligned}$$

Choose now t such that $\delta \equiv T-t$ satisfies:

$$C\big[\delta^{\gamma/2} + \delta m^{d/(d+1)}\big] = 1/2. \tag{4.23}$$

Note, for m large, that $\delta m^{d/(d+1)} \approx 1/(2C)$, and deduce in particular that the condition $r = \delta m \ge 1$ still holds in this frame. Derive that:

$$\sup_{T-\delta\le s\le T}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big] \le 2\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big] + 2C\exp(-C^{-1}m^2\delta). \tag{4.24}$$

Note now that the same strategy can be achieved on $[T-2\delta, T-\delta]$, $[T-3\delta, T-2\delta]$, $\dots$, $[T-(i+1)\delta, T-i\delta]$, $\dots$, $[0, T-N\delta]$, $N \equiv \lfloor T\delta^{-1}\rfloor$, $i+1 \le N$: due to the boundedness of u (see Theorem 3.1) and to the Hölder continuity of u (see Theorem 3.2), the restrictions of the PDE $(\mathcal{E})$ on the previous intervals fulfill Assumption **(A)** (the new final conditions fulfill **(A.1)** and **(A.4)** with respect to $\Gamma_{3.1}$, $\Gamma_{3.2}$ and α_2). In particular, (4.24) holds on each of these sets, and up to a modification, the constants C and δ can be assumed to be the same for all the intervals $[0, T-N\delta]$, $\dots$, $[T-\delta, T]$.

Hence, define:

$$\begin{aligned}a_0 &\equiv \operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_T - \bar{V}_T|^2)\big],\\ a_i &\equiv \sup_{s\in[T-i\delta, T-(i-1)\delta]}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big],\ 1 \le i \le N,\\ a_{N+1} &\equiv \sup_{s\in[0,T-N\delta]}\operatorname*{essup}_{\omega\in\Omega}\big[\Phi(|V_s - \bar{V}_s|^2)\big].\end{aligned}$$

Derive from (4.24) that:

$$\forall i \in \{0,\dots,N\},\ a_{i+1} \le 2a_i + 2C\exp(-C^{-1}m^2\delta).$$

Since a_0 reduces to zero, a discrete version of Gronwall's Lemma yields (we introduce a new constant $\bar{C}$ since C is fixed through the value of δ):

$$\begin{aligned}\forall i \in \{0,\dots,N+1\},\ a_i &\le 2C(2^i - 1)\exp(-C^{-1}m^2\delta)\\ &\le 2C\exp(i\ln(2) - C^{-1}m^2\delta)\\ &\le 2C\exp((N+1)\ln(2) - C^{-1}m^2\delta).\\ &\le \bar{C}\exp(\bar{C}T\delta^{-1} - \bar{C}^{-1}m^2\delta).\end{aligned}$$

Note, for m large enough, that $\delta m^{d/(d+1)} \geq 1/(4C)$ (cf. (4.23)), and thus, that $\delta^{-1} \leq 4Cm^{d/d+1}$. In particular, up to a modification of $\bar{C}$, we claim for m large enough:

$$\begin{aligned}\forall i \in \{0, \ldots, N+1\},\ a_i &\leq \bar{C} \exp(\bar{C}Tm^{d/(d+1)} - \bar{C}^{-1}m^2 m^{-d/(d+1)}) \\ &= \bar{C} \exp(\bar{C}Tm^{d/(d+1)} - \bar{C}^{-1}m^{(d+2)/(d+1)}).\end{aligned}$$

Take now the supremum over the indices $i \in \{0, \ldots, N+1\}$ and deduce:

$$\sup_{s \in [0,T]} \operatorname*{essup}_{\omega \in \Omega} \big[\Phi(|V_s - \overline{V}_s|^2)\big] \leq \bar{C} \exp(\bar{C}Tm^{d/(d+1)} - \bar{C}^{-1}m^{(d+2)/(d+1)}). \tag{4.25}$$

4.3.6 Conclusion

Let $m \to +\infty$ in (4.25) and derive that, for every $t \in [0,T]$, $\operatorname*{essup}_{\omega\in\Omega}[\Phi(|V_t - \bar{V}_t|^2)] = 0$. Deduce from (4.11)-(a), and from the continuity of V and $\bar{V}$ that:

$$\tilde{\mathbb{P}}\text{−a.s.},\ \forall t \in [0,T],\ V_t = \bar{V}_t = u(t, U_t). \tag{4.26}$$

Thus, from (4.11)-(b) ($\Phi' \geq c$), and (4.16), we claim:

$$\tilde{\mathbb{P}}\text{−a.s.},\ \mu\big\{t \in [0,T[,\ W_t = \bar{W}_t = \nabla_x u(t, U_t)\big\} = T, \tag{4.27}$$

where μ denotes the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

In particular $(\tilde{\Omega}, \tilde{\mathbb{P}}, \{\tilde{\mathcal{F}}\}, \tilde{B}, U)$ is a weak solution to (4.1). According to Subsection 4.2, the martingale problem associated to $(b(\cdot,\cdot,u(\cdot,\cdot),\nabla_x u(\cdot,\cdot)), a(\cdot,\cdot,u(\cdot,\cdot)))$ is well-posed. Referring to Brossard [5], we deduce that the distribution of $(U, \tilde{B})$ under $\tilde{\mathbb{P}}$ on the space $\mathcal{C}([0,T], \mathbb{R}^{2d})$ matches the law of (X, B) under $\mathbb{P}$ (the solution to (4.1) found in Subsection 4.2). Note now from Theorem 3.3 that the mapping $(\xi_s)_{s\in[0,T]} \in \mathcal{C}([0,T], \mathbb{R}^d) \mapsto (u(s,\xi_s), \nabla_x u(s,\xi_s)\mathbf{1}_{\{s<T\}})_{s\in[0,T]} \in \mathcal{C}([0,T], \mathbb{R}) \times L^2([0,T], \mathbb{R}^d)$ is continuous and thus measurable. According to (4.26) and (4.27), it is then well seen that the distribution of $(\tilde{B}, U, V, W)$ under $\tilde{\mathbb{P}}$ on the space $\mathcal{C}([0,T], \mathbb{R}^{2d+1}) \times L^2([0,T], \mathbb{R}^d)$ coincides with the distribution of $(\tilde{B}, U, \bar{V}, \bar{W})$ under $\tilde{\mathbb{P}}$ and thus with the distribution of (B, X, Y, Z) under $\mathbb{P}$. This completes the proof of the unique weak solvability of (E). □

4.3.7 Proof of Lemma 4.1

It remains to prove Lemma 4.1. We follow the proof of Lemma 1, Section 3, Chapter II in Krylov [21].

Fix $p \geq 1$ and recall from Theorem 3.5 that there exists $\gamma \in]0,1]$ (not depending on p) such that $(T-\cdot)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p$ belongs to $L^{d+1}_{\text{loc}}([0,T]\times\mathbb{R}^d, \mathbb{R})$. In particular, for a given smooth cutting function $\eta : \mathbb{R}^d \to [0,1]$, we can find a sequence $(f_n)_{n\geq 1}$ of continuous nonnegative functions with compact support such that $f_n \to (T-\cdot)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p\eta$ in $L^{d+1}([0,T]\times\mathbb{R}^d)$.

Note that the process $\bar{U}$ is, under the probability $\bar{\mathbb{P}}$, an Itô process with null drift and uniformly non-degenerate and bounded diffusion matrix. Derive in particular from Krylov's inequality (*cf.* Theorems 3 and 4, Section 3, Chapter II in Krylov [21]) that there exists a constant C, depending only on d, λ, Λ and T, such that:

$$
\begin{aligned}
&\Big|\bar{\mathbb{E}}\Big[\int_t^{\tau(t,r)} \big[(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u(s,\bar{U}_s)|^p\eta(s,\bar{U}_s)\big]ds\,\big|\tilde{\mathcal{F}}_t\Big] - \bar{\mathbb{E}}\Big[\int_t^{\tau(t,r)} f_n(s,\bar{U}_s)ds\,\big|\tilde{\mathcal{F}}_t\Big]\Big| \\
&\le \bar{\mathbb{E}}\Big[\Big(\int_t^T \big|(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u(s,\bar{U}_s)|^p\eta(s,\bar{U}_s) - f_n(s,\bar{U}_s)\big|ds\Big)\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\le C\Big[\int_t^T\int_{\mathbb{R}^d} \big|(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u(s,x)|^p\eta(s,x) - f_n(s,x)\big|^{d+1}ds\,dx\Big]^{1/(d+1)} \\
&\to 0 \text{ as } n \to +\infty.
\end{aligned}
\tag{4.28}
$$

Define now, for $m \in \mathbb{N}$, the function $K_m^{(1)} : x \in \mathbb{R} \mapsto 2^{-m}(k+1)$ for $x \in]2^{-m}k, 2^{-m}(k+1)]$, and $K_m^{(d)} : x \in \mathbb{R}^d \mapsto (K_m^{(1)}(x_1), \dots, K_m^{(1)}(x_d))$. It is readily seen that, for every $x \in \mathbb{R}^d$, $K_m^{(d)}(x) \to x$ as $m \to +\infty$. Thus, for every $n \ge 1$:

$$
\begin{aligned}
&\bar{\mathbb{E}}\Big[\int_t^{\tau(t,r)} f_n(s,\bar{U}_s)ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\le \bar{\mathbb{E}}\Big[\int_t^T f_n(s,\bar{U}_s)\mathbf{1}_{\{|\bar{U}_s-\bar{U}_t|\le r\}}ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&= \bar{\mathbb{E}}\Big[\int_t^T \lim_{m\to+\infty} f_n\big(s,\bar{U}_s-\bar{U}_t+K_m^{(d)}(\bar{U}_t)\big)\mathbf{1}_{\{|\bar{U}_s-\bar{U}_t|\le r\}}ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\le \liminf_{m\to+\infty}\bar{\mathbb{E}}\Big[\int_t^T f_n\big(s,\bar{U}_s-\bar{U}_t+K_m^{(d)}(\bar{U}_t)\big)\mathbf{1}_{\{|\bar{U}_s-\bar{U}_t|\le r\}}ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&= \liminf_{m\to+\infty}\sum_{x\in 2^{-m}\mathbb{Z}^d}\mathbf{1}_{\{K_m^{(d)}(\bar{U}_t)=x\}}\bar{\mathbb{E}}\Big[\int_t^T f_n\big(s,\bar{U}_s-\bar{U}_t+x\big)\mathbf{1}_{\{|\bar{U}_s-\bar{U}_t|\le r\}}ds\,\big|\tilde{\mathcal{F}}_t\Big].
\end{aligned}
\tag{4.29}
$$

Apply again Theorems 3 and 4, Section 3, Chapter II in Krylov [21], to the diffusion $(\bar{U}_s - \bar{U}_t)_{t\le s\le T}$ and deduce from Theorem 3.5 that for every $x \in \mathbb{R}^d$:

$$
\begin{aligned}
&\bar{\mathbb{E}}\Big[\int_t^T f_n\big(s,\bar{U}_s-\bar{U}_t+x\big)\mathbf{1}_{\{|\bar{U}_s-\bar{U}_t|\le r\}}ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\le C\Big[\int_t^T\int_{B(0,r)}\big[(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p(s,x+z)\eta(s,x+z)\big]^{d+1}ds\,dz\Big]^{1/(d+1)} \\
&\quad + C\Big[\int_t^T\int_{B(0,r)}\big|(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p(s,x+z)\eta(s,x+z) - f_n(s,x+z)\big|^{d+1}ds\,dz\Big]^{1/(d+1)} \\
&\le C\big[C_{3.5}(p(d+1))(T-t)r^d\big]^{1/(d+1)} \\
&\quad + C\Big[\int_t^T\int_{\mathbb{R}^d}\big|(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p(s,z)\eta(s,z) - f_n(s,z)\big|^{d+1}ds\,dz\Big]^{1/(d+1)}.
\end{aligned}
\tag{4.30}
$$

Thus, from (4.29) and (4.30):

$$
\begin{aligned}
&\bar{\mathbb{E}}\Big[\int_t^{\tau(t,r)} f_n(s,\bar{U}_s)ds\,\big|\tilde{\mathcal{F}}_t\Big] \\
&\le C\big[C_{3.5}(p(d+1))(T-t)r^d\big]^{1/(d+1)} \\
&\quad + C\Big[\int_t^T\int_{\mathbb{R}^d}\big|(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u|^p(s,z)\eta(s,z) - f_n(s,z)\big|^{d+1}ds\,dz\Big]^{1/(d+1)}.
\end{aligned}
\tag{4.31}
$$

Let $n \to +\infty$ and derive from (4.28) and (4.31):

$$
\Big|\bar{\mathbb{E}}\Big[\int_t^{\tau(t,r)} \big[(T-s)^{(1-\gamma)p}|\nabla^2_{x,x}u(s,\bar{U}_s)|^p\eta(s,\bar{U}_s)\big]ds\,\big|\tilde{\mathcal{F}}_t\Big]\Big| \le C\big[C_{3.5}(p(d+1))(T-t)r^d\big]^{1/(d+1)}. \tag{4.32}
$$

Let $\eta \to 1$ and complete the proof from the Beppo-Levi Theorem. □

4.4 Unique Solvability of $(\mathcal{E})$

In order to complete the proof of Theorem 2.1, we have to establish the unique solvability of the quasi-linear PDE $(\mathcal{E})$ in the space $\mathcal{V}$. To this end, note that to every $\tilde{u} \in \mathcal{V}$ and to every initial condition $(t,x) \in [0,T] \times \mathbb{R}^d$, we can associate, as done in Subsection 4.2, a weak solution $((\tilde{\Omega}, \{\tilde{\mathcal{F}}\}, \tilde{\mathbb{P}}, \tilde{B}), (\tilde{X}, \tilde{Y}, \tilde{Z}))$ to the FBSDE (E) and to the initial condition (t,x). Then Y and $\tilde{Y}$ have the same law, and thus $u(t,x) = \mathbb{E}(Y_t) = \tilde{\mathbb{E}}(\tilde{Y}_t) = \tilde{u}(t,x)$. □

5 Strategy to Estimate u

The whole sequel of the paper is devoted to the proofs of Theorems 3.3, 3.4 and 3.5. From now on, the coefficients b, f, σ and G are assumed to be smooth, *i.e.* bounded and infinitely differentiable with bounded derivatives of any order. In particular, there exists a unique bounded solution $u \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ with a bounded gradient to the quasi-linear PDE $(\mathcal{E})$. The gradient is Hölder continuous on $[0,T] \times \mathbb{R}^d$ and the partial derivatives of order two in x are also bounded and Hölder continuous, *cf.* Subsection 2.2.

Note that the generic notations "C", "C'" and "γ" denote in the sequel constants appearing in the proofs of Theorems 3.3, 3.4 and 3.5: γ always belongs to $]0,1]$. If nothing mentioned, these constants just depend on the parameters quoted in the statement of the theorem, proposition or lemma to which they refer. Of course, the values of these constants may change from line to line.

5.1 Main Tools

Our approach to estimate the derivatives of u differs from the earlier work of Delarue [8] in which the first author investigates the *supremum* norm of the gradient of u in the Lipschitz framework. The basic strategy in [8] consists in applying a variant of the Malliavin-Bismut integration by parts formula due to Thamaïer [38] (see also Fuhrman and Tessitore [14] and Thalmaïer and Wang [39]). Roughly speaking, such an integration by parts formula provides a tractable expression of the gradient of a harmonic function v associated to a given operator in terms of the values of v on a suitable boundary. This non-trivial mechanism deeply relies on the trajectories of the diffusion process χ associated to the underlying operator and in particular to the gradient of its flow. To be crude, *a priori* controls of the derivatives of χ are crucial to derive from the Malliavin-Bismut formula relevant estimates of the derivatives of the harmonic function v.

Typically, this method applies in the following way: assume that b and f vanish, choose for v the solution u and for χ the diffusion X given in (4.1), *i.e.* the solution of the SDE associated to $\sigma(\cdot,\cdot,u(\cdot,\cdot))$, and estimate the derivative of the flow of X in terms of the Lipschitz constant of σ. In our setting, even if σ is supposed to be differentiable with respect to x and y (*cf.* the beginning of Section 5), there is no hope to control efficiently the gradient of the flow of X in terms of known parameters appearing in **(A)**: under **(A)**, we just control the x-Hölder continuity of σ.

The strategy to overcome the lack of differentiability of σ under Assumption **(A)** then relies on the famous inequalities due to Calderón and Zygmund. These estimates provide a relevant L^p-control of the second order derivatives of the solution of a linear parabolic equation on $[0,T] \times \mathbb{R}^d$ with a non-degenerate and space-independent diffusion matrix, a null boundary condition and an L^p-integrable second member. From classical pertubation arguments of the diffusion matrix, similar results can be derived for operators with x-continuous second-order coefficients. For example, the Calderón and

Zygmund theory plays a crucial role in the proof of the well-posedness of the martingale problem of Stroock and Varadhan (see Stroock and Varadhan [37], Chapter VII and Appendix A).

In our setting, the ellipticity of the matrix a directly follows from Assumption **(A.2)** and its continuity in x from the *a priori* Hölder property of the solution u, see Theorem 3.2.

Our plan is then rather anachronistic: estimate first the partial derivatives of order two in x of u and then the partial derivatives of order one. To derive Theorems 3.3 and 3.4 (controls of the *supremum* and Hölder norms of $\nabla_x u$) from Theorem 3.5 (L^p-estimates of $\nabla^2_{x,x}u$), we apply again a pertubation argument by freezing the spatial parameter in the diffusion matrix $\sigma(\cdot,\cdot,u(\cdot,\cdot))$. This permits to consider u as the solution of a PDE with space-independent second-order coefficients, for which the transition densities can be explicitly written.

5.2 Pertubated Operator

Explain now in a more detailed way how to freeze the spatial parameter in the diffusion coefficient $\sigma(\cdot,\cdot,u(\cdot,\cdot))$. In short, we often write u as the solution of the following PDE:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,0,u(t,0)\big)\frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \\ \quad + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d}\big[a_{i,j}\big(t,x,u(t,x)\big) - a_{i,j}\big(t,0,u(t,0)\big)\big]\frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \\ \quad + \displaystyle\sum_{i=1}^{d} b_i\big(t,x,u(t,x),\nabla_x u(t,x)\big)\frac{\partial u}{\partial x_i}(t,x) \\ \quad + f\big(t,x,u(t,x),\nabla_x u(t,x)\big) = 0,\ (t,x)\in[0,T[\times\mathbb{R}^d, \\ u(T,x) = G(x),\ x\in\mathbb{R}^d. \end{cases} \tag{5.1}$$

Equation (5.1) expresses u as the solution of a PDE with $a(\cdot,0,u(\cdot,0))$ as diffusion matrix. Recall in this frame that the transition densities associated to the family of operators:

$$\forall t\in[0,T],\ \mathcal{L}^0_t = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}\big(t,0,u(t,0)\big)\frac{\partial^2}{\partial x_i \partial x_j}, \tag{5.2}$$

are given by the following theorem (see e.g. Section 0, Chapter VII in Stroock and Varadhan [37] for the basic result, and Corollary 4.2 and Theorem 4.6, Section 4, Chapter VI in Friedman [13] for further solvability properties of linear PDEs of parabolic type, see also Theorem 9.2.2, Section 2, Chapter IX in Krylov [25] for the Hölder regularity of the solution and of its derivatives):

Theorem 5.1 *Let $c:[0,T]\to\mathcal{S}_d(\mathbb{R})$ be a Hölder continuous function for which there exist $0<\lambda_0<\Lambda_0<+\infty$ such that $\forall t\in[0,T],\ \forall\theta\in\mathbb{R}^d,\ \lambda_0|\theta|^2 \le \langle\theta, c(t)\theta\rangle \le \Lambda_0|\theta|^2$.*

Define, $\forall 0\le t<s\le T,\ \Gamma(t,s) = \int_t^s c(u)du$, and set:

$$\forall x,y\in\mathbb{R}^d,\ \psi^{(c)}(t,x;s,y) \equiv (2\pi)^{-d/2}\big(\det[\Gamma(t,s)]\big)^{-1/2}\exp\Big[-\frac{1}{2}\langle x-y,\Gamma^{-1}(t,s)(x-y)\rangle\Big].$$

Then, for all bounded and uniformly Hölder continuous function $\varphi\in\mathcal{C}^{\beta/2,\beta}([0,T]\times\mathbb{R}^d,\mathbb{R})$, $\beta>0$, and for all bounded smooth function $h\in\mathcal{C}^{2+\beta}(\mathbb{R}^d,\mathbb{R})$ with bounded and uniformly Hölder continuous derivatives of order one and two, the function v given by:

$$\forall (t,x)\in[0,T]\times\mathbb{R}^d,\ v(t,x) = \int_{\mathbb{R}^d} h(y)\psi^{(c)}(t,x;T,y)dy + \int_t^T\int_{\mathbb{R}^d}\varphi(s,y)\psi^{(c)}(t,x;s,y)dy\,ds, \tag{5.3}$$

is the unique bounded solution in $\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$ *to the PDE:*

$$\begin{cases} \dfrac{\partial v}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} c_{i,j}(t)\dfrac{\partial^2 v}{\partial x_i \partial x_j}(t,x) + \varphi(t,x) = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d, \\ v(T,x) = h(x), \ x \in \mathbb{R}^d. \end{cases}$$

Moreover, the partial derivatives of v *of order one in* t *and of order one and two in* x *are bounded and uniformly Hölder continuous on* $[0,T]\times\mathbb{R}^d$.

The proofs of Theorems 3.3, 3.4 and 3.5 then rely on the following classical estimates of the derivatives of the kernel $\psi^{(c)}$ whose proof is left to the reader (see also Friedman [11], Chapter I, Section 4):

Proposition 5.2 *Under assumption and notation of Theorem 5.1, there exists a constant* $C_{5.2}$*, depending only on* d, λ_0 *and* Λ_0*, such that:*

$$\begin{aligned} &\forall (t,x) \in [0,T[\times\mathbb{R}^d, \ \forall (s,y) \in [0,T]\times\mathbb{R}^d, \ t<s, \\ &\quad |\nabla_x \psi^{(c)}(t,x;s,y)| \le C_{5.2}(s-t)^{-1}|x-y|\,\psi^{(c)}(t,x;s,y), \\ &\quad |\nabla^2_{x,x}\psi^{(c)}(t,x;s,y)| \le C_{5.2}(s-t)^{-1}\big[1+(s-t)^{-1}|x-y|^2\big]\psi^{(c)}(t,x;s,y). \end{aligned}$$

The following corollary is crucial in our current problem:

Corollary 5.3 *Under assumption and notation of Theorem 5.1, the first derivatives of* v *with respect to the variable* x *writes:*

$$\begin{aligned} &\forall i \in \{1,\dots,d\}, \ \forall (t,x) \in [0,T[\times\mathbb{R}^d, \\ &\frac{\partial v}{\partial x_i}(t,x) = \int_{\mathbb{R}^d} h(z)\frac{\partial \psi^{(c)}}{\partial x_i}(t,x;T,z)dz + \int_t^T \int_{\mathbb{R}^d} \varphi(s,z)\frac{\partial \psi^{(c)}}{\partial x_i}(t,x;s,z)dz\,ds. \end{aligned}$$

In particular, the following estimate holds:

$$\begin{aligned} &\forall (t,x) \in [0,T[\times\mathbb{R}^d, \\ &\quad |\nabla_x v(t,x)| \le C_{5.3}\bigg[(T-t)^{-1/2}\int_{\mathbb{R}^d}\Big[\big|h\big(x+\Gamma^{1/2}(t,T)z\big)-h(x)\big||z|\exp\big(-\frac{|z|^2}{2}\big)\Big]dz \\ &\qquad + \int_t^T (s-t)^{-1/2}\Big[\int_{\mathbb{R}^d}\big|\varphi\big(s,x+\Gamma^{1/2}(t,s)z\big)\big||z|\exp\big(-\frac{|z|^2}{2}\big)dz\Big]ds\bigg], \end{aligned} \tag{5.4}$$

where $C_{5.3}$ *is a constant that refers only to* d, λ_0 *and* Λ_0*. If* $\varphi(t,0)$ *vanishes for every* $t \in [0,T]$*, then the second order derivatives of* v *along the set* $[0,T]\times\{0\}$ *write:*

$$\begin{aligned} \forall i,j \in \{1,\dots,d\}, \ \forall t \in [0,T[, \ \frac{\partial^2 v}{\partial x_i \partial x_j}(t,0) &= \int_{\mathbb{R}^d} h(z)\frac{\partial^2 \psi^{(c)}}{\partial x_i \partial x_j}(t,0;T,z)dz \\ &\quad + \int_t^T \int_{\mathbb{R}^d} \varphi(s,z)\frac{\partial^2 \psi^{(c)}}{\partial x_i \partial x_j}(t,0;s,z)dz\,ds, \end{aligned}$$

and the following estimate holds:

$$\begin{aligned} &\forall t \in [0,T[, \\ &\quad |\nabla^2_{x,x} v(t,0)| \le C_{5.3}\bigg[(T-t)^{-1}\int_{\mathbb{R}^d}\Big[\big|h\big(\Gamma^{1/2}(t,T)z\big)-h(0)\big|\big(1+|z|^2\big)\exp\big(-\frac{|z|^2}{2}\big)\Big]dz. \\ &\qquad + \int_t^T \int_{\mathbb{R}^d}\Big[(s-t)^{-1}\big|\varphi\big(s,\Gamma^{1/2}(t,s)z\big)\big|\big(1+|z|^2\big)\exp\big(-\frac{|z|^2}{2}\big)\Big]dz\,ds\bigg]. \end{aligned} \tag{5.5}$$

Proof. Note first from Proposition 5.2 and from the Lebesgue differentiation theorem that for every $s \in]t,T]$ and every bounded and measurable function $\ell : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, the mapping:

$$L(s,.) : x \in [t,T] \times \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} \ell(s,y)\psi^{(c)}(t,x;s,y)dy,$$

is differentiable with respect to x, and that, for every $i \in \{1,\dots,d\}$, the partial derivative with respect to x_i writes:

$$\forall x \in \mathbb{R}^d,\ \frac{\partial L}{\partial x_i}(s,x) = \int_{\mathbb{R}^d} \ell(s,y)\frac{\partial \psi^{(c)}}{\partial x_i}(t,x;s,y)dy. \tag{5.6}$$

Recall now the following classical change of variables:

$$\begin{aligned}\forall (s,x) \in]t,T] \times \mathbb{R}^d,\ t<s,\ &\int_{\mathbb{R}^d} \eta(y)\psi^{(c)}(t,x;s,y)dy \\ &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} \eta\big(x+\Gamma^{1/2}(t,s)z\big)\exp\big(-\frac{|z|^2}{2}\big)dz,\end{aligned} \tag{5.7}$$

for every bounded and measurable function η from $\mathbb{R}^d$ into $\mathbb{R}$. Refer again to Proposition 5.2 and derive from (5.6) and (5.7) that:

$$\forall (s,x) \in]t,T] \times \mathbb{R}^d,\ |\nabla_x L(s,x)| \leq C(s-t)^{-1/2}\int_{\mathbb{R}^d} \big|\ell\big(x+\Gamma^{1/2}(t,s)z\big)\big||z|\exp\big(-\frac{|z|^2}{2}\big)dz.$$

Apply again the Lebesgue differentation theorem and conclude that the mapping:

$$L' : (t,x) \in [0,T] \times \mathbb{R}^d \mapsto \int_t^T \int_{\mathbb{R}^d} \ell(s,y)\psi^{(c)}(t,x;s,y)dy,$$

is differentiable with respect to x. Derive also the corresponding representation formula for $\nabla_x L'$. Apply now this procedure to h and φ and deduce from (5.3) the representation formula for the derivatives of v. Apply the same strategy to establish (5.4), but replace in addition v by $v - h(x)$, x being fixed.

The representation formula for the second order derivatives of v is proved in a similar way in the case $\varphi(\cdot,0) = 0$. Indeed, the Hölder continuity of φ ensures:

$$\big|\varphi\big(s,\Gamma^{1/2}(t,s)z\big)\big| = \big|\varphi\big(s,\Gamma^{1/2}(t,s)z\big) - \varphi(s,0)\big| \leq C'(s-t)^{\beta/2}|z|^\beta,$$

and the Lebesgue differentiation theorem still applies. □

Focus for the moment on the statement of Corollary 5.3. If the *supremum* norm of h and φ are explicitly controlled in terms of known parameters, (5.4) directly applies. However, if h or φ are just controlled in L^p, for a given $p \geq 1$, the story is rather different. Specify in this frame the meaning of h and φ in the sequel: the role of h is often played by the boundary condition G (and more generally by the solution u), which is bounded, but, φ always refers to the derivatives $\nabla^2_{x,x}u$, for which no pointwise estimate is available (*cf.* Theorem 3.5). The following lemma then gives a relevant bound of the quantities appearing in (5.4) for such a φ:

Lemma 5.4 *Keep assumption and notation of Theorem 5.1. For every $p \geq 1$, there exists a constant $C_{5.4}(p)$, depending only on d, λ_0, Λ_0 and p, such that for every $\eta \in L^p(\mathbb{R}^d)$:*

$$\begin{aligned}\forall (t,s,x) \in [0,T[\times[0,T] \times \mathbb{R}^d,\ t<s,\ &\left|\int_{\mathbb{R}^d} \eta\big(x+\Gamma^{1/2}(t,s)z\big)|z|\exp\big(-\frac{|z|^2}{2}\big)dz\right| \\ &\leq C_{5.4}(p)(s-t)^{-d/(2p)}\left[\int_{\mathbb{R}^d} |\eta|^p(z)dz\right]^{1/p}.\end{aligned}$$

Proof. Under the assumption of the statement, derive from the Hölder inequality and from the change of variables $y = x + \Gamma^{1/2}(t,s)z$:

$$\begin{aligned}
&\int_{\mathbb{R}^d} \eta\big(x + \Gamma^{1/2}(t,s)z\big)|z| \exp\big(-\frac{|z|^2}{2}\big) dz \\
&\leq C(p)\left[\int_{\mathbb{R}^d} |\eta|^p\big(x + \Gamma^{1/2}(t,s)z\big) dz\right]^{1/p} \\
&\leq C(p)(s-t)^{-d/(2p)}\left[\int_{\mathbb{R}^d} |\eta|^p(y) dy\right]^{1/p}.
\end{aligned} \tag{5.8}$$

5.3 First Example: Schauder's Theory for the Partial Derivatives of Order Two

Here is the first example of the pertubation strategy:

Theorem 5.5 *The linear PDE:*

$$\begin{cases}
\dfrac{\partial w_1}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,x,u(t,x)\big)\frac{\partial^2 w_1}{\partial x_i \partial x_j}(t,x) = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d, \\
w_1(T,x) = G(x), \ x \in \mathbb{R}^d,
\end{cases} \tag{5.9}$$

admits a unique bounded strong solution $w_1 \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d, \mathbb{R})$, with bounded and uniformly Hölder continuous derivatives of order one in t and of order one and two in x. Moreover, there exist two constants $\beta_{5.5} > 0$ and $C_{5.5}$, depending only on known parameters $\alpha_0, d, H, \lambda, \Lambda$ and T, such that:

$$\forall\ (t,x) \in [0,T[\times\mathbb{R}^d, \ |\nabla^2_{x,x} w_1(t,x)| \leq C_{5.5}(T-t)^{-1+\beta_{5.5}}. \tag{5.10}$$

Remark 5.6 This result is a specific consequence of the Schauder theory devoted to partial differential equations with Hölder continuous coefficients. We refer to Friedman [11], Chapter III, Section 2 and Chapter IV, for a complete overview of this subject, and choose to give, for the sake of completeness, the detailed proof of Theorem 5.5.

Proof (Theorem 5.5) Existence and uniqueness of a bounded strong solution to the PDE (5.9) is a direct consequence of Corollary 4.2 and Theorem 4.6, Section 4, Chapter VI in Friedman [13] (see also Theorem 9.2.2, Section 2, Chapter IX in Krylov [25] to establish the boundedness and the Hölder continuity of the derivatives up to the boundary). Turn now to (5.10) and assume w.l.o.g that x reduces to 0. Note, as done in (5.1), that the solution w_1 to (5.9) can be written as the solution of the following PDE:

$$\begin{cases}
\dfrac{\partial w_1}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,0,u(t,0)\big)\frac{\partial^2 w_1}{\partial x_i \partial x_j}(t,x) \\
\qquad + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} \big[a_{i,j}\big(t,x,u(t,x)\big) - a_{i,j}\big(t,0,u(t,0)\big)\big]\frac{\partial^2 w_1}{\partial x_i \partial x_j}(t,x) = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d, \\
w_1(T,x) = G(x), \ x \in \mathbb{R}^d.
\end{cases}$$

Deduce from Corollary 5.3 with $c(\cdot) = a(\cdot, 0, u(\cdot, 0))$ (recall that a is assumed to be smooth and is thus Hölder continuous, see the beginning of Section 5):

$$
\begin{aligned}
&|\nabla^2_{x,x} w_1(t,0)| \\
&\leq C(T-t)^{-1} \int_{\mathbb{R}^d} \Big[\big|G\big(\Gamma^{1/2}(t,T)y\big) - G(0)\big| \big[1 + |y|^2\big] \exp\big(-\frac{|y|^2}{2}\big) \Big] dy \\
&\quad + C \int_t^T \Big[\int_{\mathbb{R}^d} \Big[\big|a\big(s, \Gamma^{1/2}(t,s)y, u(s, \Gamma^{1/2}(t,s)y)\big) - a\big(s, 0, u(s,0)\big)\big| \big|\nabla^2_{x,x} w_1\big(s, \Gamma^{1/2}(t,s)y\big)\big| \\
&\qquad \times (s-t)^{-1} \big[1 + |y|^2\big] \exp\big(-\frac{|y|^2}{2}\big) \Big] dy \Big] ds.
\end{aligned}
$$

Recall now that G and $a(\cdot, \cdot, u(\cdot, \cdot))$ are Hölder continuous in x with respect to known Hölder parameters (see Assumptions **(A.3)** and **(A.4)** and Theorem 3.2). Thus, for a suitable constant $\gamma > 0$, depending on known parameters quoted in the statement of Theorem 5.5:

$$
\begin{aligned}
&|\nabla^2_{x,x} w_1(t,0)| \\
&\leq C(T-t)^{-1} \int_{\mathbb{R}^d} \Big[(T-t)^{\gamma/2} |y|^\gamma \big[1 + |y|^2\big] \exp\big(-\frac{|y|^2}{2}\big) \Big] dy \\
&\quad + C \int_t^T \Big[\int_{\mathbb{R}^d} \Big[(s-t)^{\gamma/2} |y|^\gamma \big|\nabla^2_{x,x} w_1\big(s, \Gamma^{1/2}(t,s)y\big)\big| (s-t)^{-1} \big[1 + |y|^2\big] \exp\big(-\frac{|y|^2}{2}\big) ds \Big] dy \\
&\leq C(T-t)^{-1+\gamma/2} + C \int_t^T (s-t)^{-1+\gamma/2} \sup_{y \in \mathbb{R}^d} \big[|\nabla^2_{x,x} w_1(s,y)|\big] ds.
\end{aligned}
\tag{5.11}
$$

Note first that the last term of the above inequality is finite: thanks to the regularity of a, $\nabla^2_{x,x} w_1$ is bounded on the whole set $[0,T] \times \mathbb{R}^d$. Note also that (5.11) holds in fact for every starting point $(t,x) \in [0,T[\times \mathbb{R}^d$ and deduce that:

$$
\sup_{x \in \mathbb{R}^d} |\nabla^2_{x,x} w_1(t,x)| \leq C(T-t)^{-1+\gamma/2} + C \int_t^T (s-t)^{-1+\gamma/2} \sup_{y \in \mathbb{R}^d} \big[|\nabla^2_{x,x} w_1(s,y)|\big] ds. \tag{5.12}
$$

Multiply both sides by $(T-t)^{1-\gamma/2}$ and derive that:

$$
\begin{aligned}
\sup_{x \in \mathbb{R}^d} \big[(T-t)^{1-\gamma/2} |\nabla^2_{x,x} w_1(t,x)|\big] &\leq C + C \sup_{s \in [t,T[} \sup_{y \in \mathbb{R}^d} \big[(T-s)^{1-\gamma/2} |\nabla^2_{x,x} w_1(s,y)|\big] \\
&\qquad \times (T-t)^{1-\gamma/2} \int_t^T (s-t)^{-1+\gamma/2} (T-s)^{-1+\gamma/2} ds.
\end{aligned}
\tag{5.13}
$$

Recall now that:

$$
\forall \beta_1, \beta_2 \in]0,1[, \quad \int_t^T (s-t)^{-1+\beta_1} (T-s)^{-1+\beta_2} ds = (T-t)^{-1+(\beta_1+\beta_2)} \Gamma(\beta_1)\Gamma(\beta_2)/\Gamma(\beta_1+\beta_2). \tag{5.14}
$$

Thus, from (5.13) and (5.14):

$$
\sup_{x \in \mathbb{R}^d} \big[(T-t)^{1-\gamma/2} |\nabla^2_{x,x} w_1(t,x)|\big] \leq C + C(T-t)^{\gamma/2} \sup_{s \in [t,T[} \sup_{y \in \mathbb{R}^d} \big[(T-s)^{1-\gamma/2} |\nabla^2_{x,x} w_1(s,y)|\big].
$$

Deduce for $T-t$ small enough that:

$$
\sup_{x \in \mathbb{R}^d} \big[(T-t)^{1-\gamma/2} |\nabla^2_{x,x} w_1(t,x)|\big] \leq C.
$$

Finally, following Theorems 3.1 and 3.2, we can prove that the *supremum* norm of w_1 is bounded by Λ (maximum principle for parabolic equations) and we can estimate the Hölder continuity of w_1 in terms of known parameters (Krylov and Safonov theory). In particular, for every $t \in [0,T[$, $w_1(t,\cdot)$ satisfies, up to a modification of the underlying parameters H and α_0, the same properties as G. As a consequence, the strategy applied above can be also achieved on every subinterval of $[0,T[$ of small length. This is sufficient to complete the proof. □

6 L^p-Estimates of the Second-Order Derivatives

We now investigate the second order derivatives of u in x.

6.1 Calderón and Zygmund Inequalities

The strategy relies on the Calderón and Zygmund inequalities. These estimates are well-known in the probabilistic literature since they play a fundamental role in the proof of the solvability of the martingale problem of Stroock and Varadhan. For a complete overview, we refer to Appendix A in Stroock and Varadhan [37]. We just remind the reader of the following statement:

Theorem 6.1 *Keep assumption and notation of Theorem 5.1, and assume in addition that h vanishes and that the support of the function φ is bounded. Then, the function v satisfies:*

$$\forall p \geq 1, \ \int_0^T \int_{\mathbb{R}^d} |\nabla^2_{x,x} v(t,x)|^p dt\, dx \leq C_{6.1}(p) \int_0^T \int_{\mathbb{R}^d} |\varphi(t,x)|^p dt\, dx,$$

where $C_{6.1}(p)$ just depends on d, λ_0, Λ_0 and p.

About the Proof. Note that the version stated in Stroock and Varadhan holds for $\varphi \in \mathcal{C}_0^\infty([0,T] \times \mathbb{R}^d)$. Using a standard regularization argument, derive from the Schauder theory (see e.g. Friedman [11], Chapter III, Section 2, and Chapter IV, and Krylov [25], Chapter IX) that Theorem 6.1 is still valid for $\varphi \in \mathcal{C}_0^{\beta/2,\beta}([0,T] \times \mathbb{R}^d)$, $\beta > 0$. $\square$

A commom trick consists in localizing the Calderón and Zygmund inequalities (see Gilbarg and Trudinger [15], Chapter IX, Section 5 for the original argument for elliptic equations):

Corollary 6.2 *Keep assumption and notation of Theorem 5.1, and assume in addition that h vanishes. For given $\rho > 0$ and $\theta \in]0,1[$, set $\theta' \equiv (1+\theta)/2$. Then, the function v satisfies for all $p \geq 1$ and $z \in \mathbb{R}^d$:*

$$\begin{aligned}
&(1-\theta)^{2p} \rho^{2p} \int_0^T \int_{B(z,\theta\rho)} |\nabla^2_{x,x} v(t,x)|^p dt\, dx \\
&\leq C_{6.2}(p) \left[(1-\theta')^{2p} \rho^{2p} \int_0^T \int_{B(z,\theta'\rho)} |\varphi(t,x)|^p dt\, dx + \int_0^T \int_{B(z,\theta'\rho)} |v(t,x)|^p dt\, dx \right] \\
&\quad + \frac{1}{2} (1-\theta')^{2p} \rho^{2p} \int_0^T \int_{B(z,\theta'\rho)} |\nabla^2_{x,x} v(t,x)|^p dt\, dx,
\end{aligned}$$

where $C_{6.2}(p)$ depends only on d, λ_0, Λ_0 and p.

Proof. Fix $\theta \in]0,1[$ and set $\theta' \equiv (1+\theta)/2$. Assume without loss of generality that $z = 0$ and focus on the product $\tilde{v} : (t,x) \in [0,T] \times \mathbb{R}^d \mapsto \eta(x) v(t,x)$ for a given cutting function $\eta : \mathbb{R}^d \to [0,1]$ such that:

$$\begin{cases}
\forall x \in \mathbb{R}^d, \ |x| \leq \theta\rho, \ \eta(x) = 1, \\
\forall x \in \mathbb{R}^d, \ |x| \geq \theta'\rho, \ \eta(x) = 0, \\
\forall x \in \mathbb{R}^d, \ \theta\rho \leq |x| \leq \theta'\rho, \ |\nabla_x \eta(x)| \leq k(1-\theta)^{-1}\rho^{-1}, \\
\forall x \in \mathbb{R}^d, \ \theta\rho \leq |x| \leq \theta'\rho, \ |\nabla^2_{x,x} \eta(x)| \leq k(1-\theta)^{-2}\rho^{-2},
\end{cases} \tag{6.1}$$

for a suitable constant k (not depending on θ). Note that $\tilde{v}$ satisfies the PDE:

$$\begin{aligned}
&\frac{\partial \tilde{v}}{\partial t}(t,x) + \frac{1}{2} \sum_{i,j=1}^d c_{i,j}(t) \frac{\partial^2 \tilde{v}}{\partial x_i \partial x_j}(t,x) \\
&= -\eta(x)\varphi(t,x) + \frac{1}{2} \sum_{i,j=1}^d \left[c_{i,j}(t) \frac{\partial^2 \eta}{\partial x_i \partial x_j}(x) v(t,x) \right] + \langle \nabla_x \eta(x), c(t) \nabla_x v(t,x) \rangle \\
&\equiv -\tilde{\varphi}(t,x), \ (t,x) \in [0,T[\times \mathbb{R}^d.
\end{aligned} \tag{6.2}$$

Derive in particular from Theorem 6.1:

$$\begin{aligned}
&\int_0^T \int_{B(0,\theta\rho)} |\nabla_{x,x} v(t,x)|^p dt\, dx \\
&\quad \leq \int_0^T \int_{B(0,\theta'\rho)} |\nabla_{x,x} \tilde{v}(t,x)|^p dt\, dx \\
&\quad \leq C \int_0^T \int_{B(0,\theta'\rho)} |\tilde{\varphi}(t,x)|^p dt\, dx \\
&\quad \leq C \int_0^T \int_{B(0,\theta'\rho)} \big[\eta^p(x) |\varphi(t,x)|^p\big] dt\, dx \\
&\qquad + C \int_0^T \int_{B(0,\theta'\rho)} \big[|v(t,x)|^p |\nabla^2_{x,x}\eta(x)|^p + |\nabla_x v(t,x)|^p |\nabla_x \eta(x)|^p\big] dt\, dx.
\end{aligned}$$

Thanks to (6.1), derive that:

$$\begin{aligned}
\int_0^T \int_{B(0,\theta\rho)} |\nabla_{x,x} v(t,x)|^p dt\, dx \leq C \int_0^T \int_{B(0,\theta'\rho)} |\varphi(t,x)|^p dt\, dx \\
+ C(1-\theta)^{-2p}\rho^{-2p} \int_0^T \int_{B(0,\theta'\rho)} |v(t,x)|^p dt\, dx \\
+ C(1-\theta)^{-p}\rho^{-p} \int_0^T \int_{B(0,\theta'\rho)} |\nabla_x v(t,x)|^p dt\, dx.
\end{aligned} \tag{6.3}$$

Recall now the following Gagliardo-Nirenberg inequality (see e.g. Theorem 10.1, page 27, in Friedman [12], see also page 125 in Nirenberg [31], and Theorem 7.28, Chapter VII, in Gilbarg and Trudinger [15]):

Lemma 6.3 *Let $q_1, q_2 \in [1, +\infty]$. Assume that p^{-1} writes $p^{-1} = (2q_1)^{-1} + (2q_2)^{-1}$. Then, there exists a constant $C_{6.3}(p, q_1, q_2)$, depending only on p, q_1 and q_2, such that for every smooth function ℓ from $\overline{B(0,1)}$ into $\mathbb{R}$:*

$$\begin{aligned}
\int_{B(0,1)} |\nabla_x \ell(x)|^p dx \leq C_{6.3}(p, q_1, q_2) \bigg[\int_{B(0,1)} \big[|\nabla^2_{x,x}\ell(x)|^{q_1} + |\nabla_x \ell(x)|^{q_1} + |\ell(x)|^{q_1}\big] dx\bigg]^{p/(2q_1)} \\
\times \bigg[\int_{B(0,1)} |\ell(x)|^{q_2} dx\bigg]^{p/(2q_2)}.
\end{aligned} \tag{6.4}$$

Apply Lemma 6.3 with $q_1 = q_2 = p$. Note then from the Young inequality and from a scaling argument that for every $r > 0$ and for every smooth function $\ell : \overline{B(0,r)} \to \mathbb{R}$:

$$\forall \varepsilon > 0, \int_{B(0,r)} |\nabla_x \ell(x)|^p dx \leq \varepsilon r^p \int_{B(0,r)} |\nabla^2_{x,x}\ell(x)|^p dx + C(p)(1+\varepsilon^{-1}) r^{-p} \int_{B(0,r)} |\ell(x)|^p dx, \tag{6.5}$$

where $C(p)$ refers only to p.

Multiply (6.3) by $(1-\theta)^{2p}\rho^{2p}$ and apply (6.5) with $\ell = v(t,\cdot)$, $\varepsilon = (2^{2p+1}C)^{-1}(1-\theta)^p$ and $r = \theta'\rho$

(note that $r^{-p} \leq 2^p \rho^{-p}$ since $\theta' \geq 1/2$):

$$
\begin{aligned}
&(1-\theta)^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta\rho)} |\nabla_{x,x} v(t,x)|^p dt\,dx \\
&\leq C(1-\theta)^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta'\rho)} |\varphi(t,x)|^p dt\,dx \\
&\quad + C\int_0^T\int_{B(0,\theta'\rho)} |v(t,x)|^p dt\,dx + C(1-\theta)^p\rho^p\int_0^T\int_{B(0,\theta'\rho)} |\nabla_x v(t,x)|^p dt\,dx \\
&\leq C(1-\theta)^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta'\rho)} |\varphi(t,x)|^p dt\,dx \\
&\quad + C'\int_0^T\int_{B(0,\theta'\rho)} |v(t,x)|^p dt\,dx + 2^{-(2p+1)}(1-\theta)^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta'\rho)} |\nabla^2_{x,x} v(t,x)|^p dt\,dx.
\end{aligned}
$$

Note that $1-\theta' = (1-\theta)/2$. Thus:

$$
\begin{aligned}
&(1-\theta)^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta\rho)} |\nabla_{x,x} v(t,x)|^p dt\,dx \\
&\leq C'(1-\theta')^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta'\rho)} |\varphi(t,x)|^p dt\,dx \qquad (6.6)\\
&\quad + C'\int_0^T\int_{B(0,\theta'\rho)} |v(t,x)|^p dt\,dx + \frac{1}{2}(1-\theta')^{2p}\rho^{2p}\int_0^T\int_{B(0,\theta'\rho)} |\nabla^2_{x,x} v(t,x)|^p dt\,dx.
\end{aligned}
$$

This completes the proof. □

6.2 Proof of Theorem 3.5

We now prove Theorem 3.5. We start with the following property:

Theorem 6.4 *There exists a constant $\beta_{6.4} \in]0,1]$, depending only on $\alpha_0, d, H, \lambda, \Lambda$ and T, such that:*

$$
\begin{aligned}
\forall p \geq 1,\ R \geq 1,\ \delta \in]0,T],\ z \in \mathbb{R}^d,\ &\int_{T-\delta}^T\int_{B(z,R)} \big[(T-s)^{1-\beta_{6.4}}\big(|\nabla_x u(s,y)|^2 + |\nabla^2_{x,x} u(s,y)|\big)\big]^p ds\,dy \\
&\leq C_{6.4}(p)\delta R^d,
\end{aligned}
$$

where $C_{6.4}(p)$ depends only on $\alpha_0, d, H, \lambda, \Lambda, p$ and T.

Proof of Theorem 6.4. Fix $p \geq 1$ and assume w.l.o.g that $z = 0$. Consider first the following linear parabolic equation:

$$
\begin{cases}
\partial_t w_2(t,x) + \mathcal{L}_t w_2(t,x) = -g(t,x),\ (t,x) \in [0,T[\times\mathbb{R}^d, \\
w_2(T,x) = 0,\ x \in \mathbb{R}^d,
\end{cases} \tag{6.7}
$$

where $\mathcal{L}_t$ denotes the second-order differential operator:

$$
\mathcal{L}_t = \frac{1}{2}\sum_{i,j=1}^d a_{i,j}\big(t,x,u(t,x)\big)\frac{\partial^2}{\partial x_i \partial x_j}, \tag{6.8}
$$

and g the non-linear term:

$$
\forall (t,x) \in [0,T]\times\mathbb{R}^d,\ g(t,x) = f\big(t,x,u(t,x),\nabla_x u(t,x)\big) + \big\langle b\big(t,x,u(t,x),\nabla_x u(t,x)\big), \nabla_x u(t,x)\big\rangle. \tag{6.9}
$$

Recall then that the coefficients $a(t,x,u(t,x))$ and $g(t,x)$ are Hölder continuous with respect to (t,x) (see Subsection 2.2). Deduce in particular without any difficulties that (6.7) admits a unique bounded solution $w_2 \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ with bounded and uniformly Hölder continuous partial derivatives of order one in t and of order one and two in x (see again Friedman [13], Chapter VI, Section 4, Corollary 4.2 and Theorem 4.6, and Krylov [25], Chapter IX, Section 2, Theorem 9.2.2).

Note that the function $u - w_2$ matches the solution w_1 to the PDE (5.9) (*cf.* Theorem 5.5). Recall that the *supremum* and Hölder norms of w_1 can be estimated in terms of known parameters. Thus, referring to Theorem 5.5 (local boundedness of the derivatives of order two of w_1), there exist constants $\beta > 0$ and C (depending on parameters quoted in Theorem 3.5), such that:

$$\begin{aligned}
&\forall (t,x) \in [0,T[\times\mathbb{R}^d, \begin{cases} |w_1(t,x)| \le C, \\ |\nabla^2_{x,x} w_1(t,x)| \le C(T-t)^{-(1-\beta/2)}, \end{cases} \\
&\forall (t,x),\ (t',x') \in [0,T[\times\mathbb{R}^d,\ |w_1(t,x) - w_1(t',x')| \le C\big[|t-t'|^{\beta/2} + |x-x'|^{\beta}\big].
\end{aligned} \tag{6.10}$$

Hence, for $0 < \gamma \le \beta/2$:

$$\forall R > 0,\ \forall \delta \in]0,T],\ \int_{T-\delta}^{T}\int_{B(0,R)} \big[(T-s)^{1-\gamma}|\nabla^2_{x,x} w_1(s,y)|\big]^p ds\, dy \le C(p)\delta R^d. \tag{6.11}$$

Thus, it remains to focus on the second-order derivatives of w_2. Multiply $w_2(t,x)$ by $(T-t)^{1-\gamma}$ and derive that:

$$\frac{\partial}{\partial t}\big[(T-t)^{1-\gamma} w_2(t,x)\big] + \mathcal{L}_t\big[(T-t)^{1-\gamma} w_2(t,x)\big] = -(T-t)^{1-\gamma} g(t,x) + (1-\gamma)(T-t)^{-\gamma} w_2(t,x).$$

Thanks to Theorems 3.1 and 3.2 and to (6.10), w_2 is bounded and uniformly Hölder continuous on $[0,T]\times\mathbb{R}^d$. Thus, for $\gamma \le (\beta \wedge \alpha_2)/2$:

$$\forall (t,x) \in [0,T[\times\mathbb{R}^d,\ (T-t)^{-\gamma}|w_2(t,x)| = (T-t)^{-\gamma}|w_2(t,x) - w_2(T,x)| \le C. \tag{6.12}$$

Recall (*cf.* (5.2)) that $\mathcal{L}^0_t = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,0,u(t,0))\frac{\partial^2}{\partial x_i \partial x_j}$ and write w_2 as the solution of:

$$\begin{aligned}
&\frac{\partial}{\partial t}\big[(T-t)^{1-\gamma} w_2(t,x)\big] + \mathcal{L}^0_t\big[(T-t)^{1-\gamma} w_2(t,x)\big] \\
&\quad = -(T-t)^{1-\gamma} g(t,x) + \big[\mathcal{L}^0_t - \mathcal{L}_t\big]\big[(T-t)^{1-\gamma} w_2(t,x)\big] + (1-\gamma)(T-t)^{-\gamma} w_2(t,x).
\end{aligned}$$

The function g and the partial derivatives of order two of w_2 are Hölder continuous. Note also that $(T-t)^{-\gamma} w_2(t,x)$ writes

$$(T-t)^{-\gamma} w_2(t,x) = -(T-t)^{-\gamma}\int_t^T \frac{\partial w_2}{\partial t}(s,x) ds.$$

Since the coefficients of (6.7) are differentiable with respect to x, with bounded and Hölder continuous derivatives, it is well seen that $\partial w_2/\partial t$ is bounded and differentiable with respect to x, with bounded derivatives. In particular, the function $(T-\cdot)^{-\gamma} w_2$ is Hölder continuous.

Hence, we can apply Corollary 6.2 on the interval $]T-\delta, T[$ to $v = (T-\cdot)^{1-\gamma} w_2$ and $c = a(\cdot, 0, u(\cdot, 0))$.

Derive for all $\rho > 0$, $\theta \in]0,1[$ and $\theta' = (1+\theta)/2$:

$$\begin{aligned}
&(1-\theta)^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta\rho)}(T-t)^{(1-\gamma)p}\big|\nabla^2_{x,x}w_2(t,x)\big|^p dt\,dx \\
&\quad\le C(1-\theta')^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|g(t,x)|^p dt\,dx \\
&\qquad + C(1-\theta')^{2p}\rho^{2p}\sum_{i,j=1}^{d}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}\Big[(T-t)^{(1-\gamma)p} \\
&\qquad\quad \times \big|a_{i,j}(t,x,u(t,x)) - a_{i,j}(t,0,u(t,0))\big|^p\big|\nabla^2_{x,x}w_2(t,x)\big|^p\Big]dt\,dx \\
&\qquad + C(1-\theta')^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{-\gamma p}|w_2(t,x)|^p dt\,dx \\
&\qquad + C\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|w_2(t,x)|^p dt\,dx \\
&\qquad + \frac{1}{2}(1-\theta')^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla^2_{x,x}w_2(t,x)|^p dt\,dx \\
&\quad\equiv T(1)+T(2)+T(3)+T(4)+T(5).
\end{aligned} \tag{6.13}$$

Recall first that the mapping $a(\cdot,\cdot,u(\cdot,\cdot))$ is uniformly Hölder continuous in x with exponent α_2 (see Assumptions **(A.3)** and **(A.4)** and Theorem 3.2). Thus,

$$T(2) \le C(1-\theta')^{2p}\rho^{(2+\alpha_2)p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}\big|\nabla^2_{x,x}w_2(t,x)\big|^p dt\,dx. \tag{6.14}$$

Deduce from (6.12) that:

$$T(3) \le C\delta(1-\theta')^{2p}\rho^{2p+d},\ T(4) \le C(p)\delta\rho^d. \tag{6.15}$$

Turn finally to $T(1)$ and note to this end from Assumption **(A.1)** (growth of b and f), from Theorem 3.1 (boundedness of u) and from (6.9) (definition of g) that:

$$T(1) \le C(1-\theta')^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}\big(1+|\nabla_x u(t,x)|^{2p}\big)dt\,dx. \tag{6.16}$$

Apply now Lemma 6.3 to the triple $(2p, p, +\infty)$. Deduce from the Young inequality and from a scaling argument that for every smooth function $\ell : \overline{B(0,r)} \to \mathbb{R}$, $r > 0$:

$$\begin{aligned}
&\int_{B(0,r)}|\nabla_x \ell(x)|^{2p}dx \\
&\quad\le C(p)\sup_{x\in B(0,r)}\big[|\ell(x)|^p\big]\Big[\int_{B(0,r)}|\nabla^2_{x,x}\ell(x)|^p dx + r^{-2p}(1+r^d)\sup_{x\in B(0,r)}\big[|\ell(x)|^p\big]\Big],
\end{aligned} \tag{6.17}$$

where $C(p)$ refers only to p. Apply (6.17) to $\ell = u(t,\cdot) - u(t,0)$ and deduce from the Hölder continuity of u that:

$$\int_{B(0,\theta'\rho)}|\nabla_x u(t,x)|^{2p}dx \le C\rho^{\gamma p}\Big[\int_{B(0,\theta'\rho)}|\nabla^2_{x,x}u(t,x)|^p dx + \rho^{-2p}(1+\rho^d)\Big]. \tag{6.18}$$

Assume from now on that $\rho \le 1$. Derive then from (6.16) and (6.18) that:

$$T(1) \le C(1-\theta')^{2p}\rho^{(2+\gamma)p}\Big[\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla^2_{x,x}u(t,x)|^p dt\,dx + \delta\rho^{-2p}\Big]. \tag{6.19}$$

Deduce finally from (6.13), (6.14), (6.15) and (6.19):

$$
\begin{aligned}
&(1-\theta)^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta\rho)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 w_2(t,x)\big|^p dt\,dx\\
&\quad\le (1-\theta')^{2p}\rho^{2p}(C\rho^{\gamma p}+1/2)\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla_{x,x}^2 w_2(t,x)|^p dt\,dx\\
&\qquad + C(1-\theta')^{2p}\rho^{(2+\gamma)p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla_{x,x}^2 u(t,x)|^p dt\,dx + C\delta.
\end{aligned}
$$

Referring to (6.11), note that the above expression still holds with $\nabla_{x,x}^2 u$ instead of $\nabla_{x,x}^2 w_2$:

$$
\begin{aligned}
&(1-\theta)^{2p}\rho^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta\rho)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 u(t,x)\big|^p dt\,dx\\
&\quad\le (1-\theta')^{2p}\rho^{2p}(C\rho^{\gamma p}+1/2)\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla_{x,x}^2 u(t,x)|^p dt\,dx + C\delta.
\end{aligned}
$$

Deduce that:

$$
\begin{aligned}
&\rho^{2p}\sup_{\theta\in]0,1[}\left[(1-\theta)^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta\rho)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 u(t,x)\big|^p dt\,dx\right]\\
&\quad\le \rho^{2p}(C\rho^{\gamma p}+1/2)\sup_{\theta'\in]0,1[}\left[(1-\theta')^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta'\rho)}(T-t)^{(1-\gamma)p}|\nabla_{x,x}^2 u(t,x)|^p dt\,dx\right] + C\delta.
\end{aligned}
$$

Choose $\rho=\rho_1\equiv\min(\rho_0,1)$, with $C\rho_0^{\gamma p}+1/2=3/4$, and derive that:

$$
\rho_1^{2p}\sup_{\theta\in]0,1[}\left[(1-\theta)^{2p}\int_{T-\delta}^{T}\int_{B(0,\theta\rho_1)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 u(t,x)\big|^p dt\,dx\right]\le 4C\delta.
$$

Thus:

$$
\int_{T-\delta}^{T}\int_{B(0,\rho_1/2)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 u(t,x)\big|^p dt\,dx \le 4C\delta\rho_1^{-2p}\equiv C'\delta. \tag{6.20}
$$

Note in fact that (6.20) holds for every ball $B(z,\rho_1/2)$, $z\in\mathbb{R}^d$. Choose then $R\ge 1$. Since the ball $B(0,R)$ can be covered by $N\times\lfloor R\rho_1^{-1}\rfloor^d$ balls of radius $\rho_1/2$ (for a suitable universal integer $N\ge 1$), deduce:

$$
\int_{T-\delta}^{T}\int_{B(0,R)}(T-t)^{(1-\gamma)p}\big|\nabla_{x,x}^2 u(t,x)\big|^p dt\,dx \le C'\delta R^d.
$$

Apply again (6.17) and complete the proof of Theorem 6.4. $\square$

The complete statement of Theorem 3.5 (estimate of $\partial_t u$) then follows from Theorem 6.4 and from the growth properties of b, f and σ (*cf.* Assumption **(A.1)**).

7 Pointwise Estimates of the Gradient

7.1 Proof of Theorem 3.3

We now turn to the proof of Theorem 3.3. For the sake of simplicity, we assume that x reduces to 0: we thus aim to control the quantity $\nabla_x u(t,0)$, for $t\in[0,T[$. To this end, we follow again the pertubation argument exposed in Section 5, but, introduce in addition a localization procedure. Denote indeed by $\eta\in\mathcal{C}^\infty(\mathbb{R}^d,[0,1])$ a smooth function matching 1 on the ball $B(0,1)$ and vanishing outside the ball $B(0,2)$. The function $\tilde{u} \equiv u\eta$ satisfies the following PDE (*cf.* (6.9) for the definition of g):

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,0,u(t,0)\big)\dfrac{\partial^2 \tilde{u}}{\partial x_i \partial x_j}(t,x) \\ \quad +\dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d} \big[a_{i,j}\big(t,x,u(t,x)\big) - a_{i,j}\big(t,0,u(t,0)\big)\big]\dfrac{\partial^2 \tilde{u}}{\partial x_i \partial x_j}(t,x) \\ \quad +\eta(x) g\big(t,x,u(t,x),\nabla_x u(t,x)\big) - \langle \nabla_x \eta(x), a\big(t,x,u(t,x)\big)\nabla_x u(t,x)\rangle \\ \quad -\dfrac{1}{2}u(t,x)\displaystyle\sum_{i,j=1}^{d} a_{i,j}\big(t,x,u(t,x)\big)\dfrac{\partial^2 \eta}{\partial x_i \partial x_j}(x) = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d, \\ \tilde{u}(T,x) = G(x)\eta(x), \ x \in \mathbb{R}^d. \end{cases} \tag{7.1}$$

Define for the sake of simplicity:

$$\begin{aligned} \tilde{g}(t,x) \equiv\ & \frac{1}{2}\sum_{i,j=1}^{d} \big[a_{i,j}\big(t,x,u(t,x)\big) - a_{i,j}\big(t,0,u(t,0)\big)\big]\frac{\partial^2 \tilde{u}}{\partial x_i \partial x_j}(t,x) \\ & + \eta(x) g\big(t,x,u(t,x),\nabla_x u(t,x)\big) - \langle \nabla_x \eta(x), a\big(t,x,u(t,x)\big)\nabla_x u(t,x)\rangle \\ & - \frac{1}{2}u(t,x)\sum_{i,j=1}^{d} a_{i,j}\big(t,x,u(t,x)\big)\frac{\partial^2 \eta}{\partial x_i \partial x_j}(x), \\ \tilde{G}(x) \equiv\ & G(x)\eta(x). \end{aligned} \tag{7.2}$$

Derive from Assumption **(A.1)** (growth of the coefficients) and Theorem 3.1 (boundedness of u) that:

$$\forall (t,x) \in [0,T[\times\mathbb{R}^d, \ |\tilde{g}(t,x)| \le C\big(1 + |\nabla_x u(t,x)|^2 + |\nabla_{x,x} u(t,x)|\big)\mathbf{1}_{\{|x|\le 2\}} \tag{7.3}$$

Apply now Corollary 5.3 to $c(\cdot) = a(\cdot,0,u(\cdot,0))$ and deduce for every $t \in [0,T[$:

$$\begin{aligned} \big|\nabla_x u(t,0)\big| &= \big|\nabla_x \tilde{u}(t,0)\big| \\ &\le C\bigg[(T-t)^{-1/2}\int_{\mathbb{R}^d}\bigg[\big|\tilde{G}\big(\Gamma^{1/2}(t,T)z\big) - \tilde{G}(0)\big||z|\exp\big(-\frac{|z|^2}{2}\big)\bigg]dz \\ &\qquad + \int_t^T (s-t)^{-1/2}\bigg[\int_{\mathbb{R}^d}\big|\tilde{g}\big(s,\Gamma^{1/2}(t,s)z\big)\big||z|\exp\big(-\frac{|z|^2}{2}\big)dz\bigg]ds\bigg] \\ &\equiv T(1) + T(2). \end{aligned} \tag{7.4}$$

Note now from Assumption **(A.1)** (growth of σ and G), from the boundedness of u (see Theorem 3.1) and from Assumption **(A.4)** (Hölder continuity of G) that:

$$T(1) \le C(T-t)^{-(1-\alpha_0)/2}. \tag{7.5}$$

Turn now to $T(2)$. Set $\gamma = \beta_{6.4}$ (*cf.* Theorem 6.4) and consider a small real $\varepsilon > 0$. Deduce from Lemma 5.4 (with $p = d/\varepsilon$) and from the Hölder inequality that there exists a constant $C(\varepsilon)$, depending only on ε and on known parameters quoted in the statement of Theorem 3.3, such that:

$$\begin{aligned}
T(2) &\le C(\varepsilon) \int_t^T \left[(s-t)^{-(1+\varepsilon)/2} \left[\int_{B(0,2)} |\tilde{g}(s,z)|^{d/\varepsilon} dz \right]^{\varepsilon/d} \right] ds \\
&= C(\varepsilon) \int_t^T \left[(s-t)^{-(1+\varepsilon)/2} (T-s)^{-(1-\gamma)} \left[\int_{B(0,2)} (T-s)^{d(1-\gamma)/\varepsilon} |\tilde{g}(s,z)|^{d/\varepsilon} dz \right]^{\varepsilon/d} \right] ds \\
&\le C(\varepsilon) \left[\int_t^T (s-t)^{-d(1+\varepsilon)/(2(d-\varepsilon))} (T-s)^{-d(1-\gamma)/(d-\varepsilon)} \, ds \right]^{1-\varepsilon/d} \\
&\quad \times \left[\int_t^T \int_{B(0,2)} (T-s)^{d(1-\gamma)/\varepsilon} |\tilde{g}(s,z)|^{d/\varepsilon} dz \, ds \right]^{\varepsilon/d}.
\end{aligned} \tag{7.6}$$

Deduce finally from (5.14), (7.3) and Theorem 6.4 that for ε small enough :

$$\begin{aligned}
T(2) &\le C(\varepsilon)(T-t)^{1-\varepsilon/d-(1+\varepsilon)/2-(1-\gamma)} (T-t)^{\varepsilon/d} \\
&= C(\varepsilon)(T-t)^{-(1+\varepsilon)/2+\gamma}.
\end{aligned} \tag{7.7}$$

Choose ε small enough so that:

$$T(2) \le C(T-t)^{-(1-\gamma)/2}. \tag{7.8}$$

From (7.4), (7.5) and (7.8), we complete the proof. □

7.2 Proof of Theorem 3.4

We finally prove the Hölder estimate of the gradient. We first investigate the regularity of $\nabla_x u$ with respect to the variable x.

Lemma 7.1 *There exist two constants $\beta_{7.1} > 0$ and $C_{7.1}$, depending only on known parameters quoted in Theorem 3.4, such that for every $t \in [0,T[$, for every $(x,y) \in \mathbb{R}^d$:*

$$|\nabla_x u(t,x) - \nabla_x u(t,y)| \le C_{7.1}(T-t)^{-(1-\beta_{7.1})/2} |x-y|^{\beta_{7.1}}.$$

Proof. Assume w.l.o.g that $y = 0$. Note from Theorem 3.3 (local boundedness of the gradient) that we can also assume that $|x| \le 1$. This permits to truncate the function u as done in the latter subsection: multiply u by a smooth cutting function $\eta : \mathbb{R}^d \to [0,1]$, matching 1 on $B(0,1)$ and vanishing outside $B(0,2)$. Recall then that the product $\tilde{u} \equiv u\eta$ satisfies the following PDE:

$$\begin{cases}
\dfrac{\partial \tilde{u}}{\partial t}(t,x) + \dfrac{1}{2} \displaystyle\sum_{i,j=1}^d a_{i,j}\big(t,0,u(t,0)\big) \dfrac{\partial^2 \tilde{u}}{\partial x_i \partial x_j}(t,x) + \tilde{g}(t,x) = 0, \ (t,x) \in [0,T[\times \mathbb{R}^d, \\
\tilde{u}(T,x) = \tilde{G}(x), \ x \in \mathbb{R}^d,
\end{cases} \tag{7.9}$$

where $\tilde{g}$ and $\tilde{G}$ are given by (7.2) and satisfy (7.3). According to Theorem 5.1 and to Corollary 5.3 (with $c(\cdot) = a(\cdot,0,u(\cdot,0))$), the partial derivative $\partial \tilde{u}/\partial x_i(t,x)$, for a given $i \in \{1,\dots,d\}$, writes for every $(t,x) \in [0,T[\times \mathbb{R}^d$:

$$\begin{aligned}
\frac{\partial \tilde{u}}{\partial x_i}(t,x) &= \int_{\mathbb{R}^d} \tilde{G}(z) \frac{\partial \psi^{(c)}}{\partial x_i}(t,x;T,z) dz + \int_t^T \int_{\mathbb{R}^d} \tilde{g}(s,y) \frac{\partial \psi^{(c)}}{\partial x_i}(t,x;s,y) dy \, ds \\
&\equiv R(1,x) + R(2,x).
\end{aligned} \tag{7.10}$$

Note that $x \in \mathbb{R}^d \mapsto R(1,x)$ is differentiable with respect to x. Using the Hölder continuity of the boundary condition $\tilde{G}$ (*cf.* Assumptions **(A.1)** and **(A.4)**), we can show as in the proof of Theorem

5.5, see (5.11) (the proof is even easier in this new case since the operator is space-independent), that:

$$\forall x \in \mathbb{R}^d,\ |\nabla_x R(1,x)| \leq C(T-t)^{-(1-\gamma)},$$

for a suitable constant $\gamma > 0$. Moreover, following the proof of Theorem 3.3 (Estimate of $T(1)$, *cf.* (7.5)), it is well-seen that:

$$\forall x \in \mathbb{R}^d,\ |R(1,x)| \leq C(T-t)^{-(1-\gamma)/2}.$$

Thus, for every $\varepsilon \in]0,1]$:

$$\begin{aligned}\forall x \in \mathbb{R}^d,\ |R(1,x) - R(1,0)| &\leq 2 \sup_{y\in\mathbb{R}^d} |R(1,y)|^{1-\varepsilon} \sup_{y\in\mathbb{R}^d} |\nabla_x R(1,y)|^{\varepsilon} |x|^{\varepsilon} \\ &\leq C(T-t)^{-(1-\gamma)(1-\varepsilon)/2-(1-\gamma)\varepsilon}|x|^{\varepsilon} \\ &= C(T-t)^{-(1-\gamma)(1/2+\varepsilon/2)}|x|^{\varepsilon}.\end{aligned}$$

Hence, for a new $\gamma > 0$:

$$\forall x \in \overline{B(0,1)},\ |R(1,x) - R(1,0)| \leq C(T-t)^{-(1-\gamma/2)/2}|x|^{\gamma}. \tag{7.11}$$

Turn now to $R(2,x)$. Note that $R(2,x)$ writes:

$$R(2,x) \equiv \int_t^T r(2,s,x)ds,\ r(2,s,x) \equiv \int_{\mathbb{R}^d} \tilde{g}(s,y)\frac{\partial \psi^{(c)}}{\partial x_i}(t,x;s,y)dy. \tag{7.12}$$

For every $s \in]t,T[$, the function $r(2,s,\cdot)$ is differentiable with respect to x. Following Corollary 5.3 and Lemma 5.4, for every $\varepsilon \in]0,1]$, there exists a constant $C(\varepsilon)$, depending only on ε and on known parameters quoted in the statement of Theorem 3.4, such that:

$$\begin{aligned}|\nabla_x r(2,s,x)| &\leq C(\varepsilon)(s-t)^{-1}\int_{\mathbb{R}^d} |\tilde{g}(s,x+z\Gamma^{1/2}(t,s))|(1+|z|^2)\exp\big(-\frac{|z|^2}{2}\big)dz \\ &\leq C(\varepsilon)(s-t)^{-(1+\varepsilon/2)}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d}.\end{aligned} \tag{7.13}$$

Derive from (7.13) that:

$$\forall x \in \mathbb{R}^d,\ |r(2,s,x) - r(2,s,0)| \leq C(\varepsilon)|x|(s-t)^{-(1+\varepsilon/2)}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d}.$$

Note again from Corollary 5.3 and Lemma 5.4 that:

$$\forall x \in \mathbb{R}^d,\ |r(2,s,x)| \leq C(\varepsilon)(s-t)^{-(1+\varepsilon)/2}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d}.$$

Derive finally that:

$$\forall x \in \mathbb{R}^d,\ |r(2,s,x) - r(2,s,0)| \leq C(\varepsilon)|x|^{\varepsilon}(s-t)^{-\varepsilon(1+\varepsilon/2)-(1+\varepsilon)(1-\varepsilon)/2}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d}.$$

Thus,

$$\forall x \in \mathbb{R}^d,\ |r(2,s,x) - r(2,s,0)| \leq C(\varepsilon)|x|^{\varepsilon}(s-t)^{-1/2-\varepsilon}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d}. \tag{7.14}$$

Derive from (7.12) and (7.14) that for every $x \in \mathbb{R}^d$:

$$|R(2,x) - R(2,0)| \leq C(\varepsilon)|x|^{\varepsilon}\int_t^T (s-t)^{-1/2-\varepsilon}\left[\int_{\mathbb{R}^d} |\tilde{g}(s,z)|^{d/\varepsilon}dz\right]^{\varepsilon/d} ds.$$

Follow (7.6) and (7.7) in the proof of Theorem 3.3 and deduce for a suitable $\gamma > 0$:

$$\forall x \in \overline{B(0,1)},\ |R(2,x) - R(2,0)| \le C(T-t)^{-(1-\gamma/2)/2}|x|^{\gamma}. \tag{7.15}$$

Thanks to (7.10), (7.11) and (7.15), we complete the proof. □

We are now in position to complete the proof of Theorem 3.4. It is sufficient to prove the following Lemma:

Lemma 7.2 *There exist two constants $\beta_{7.2} > 0$ and $C_{7.2}$, depending only on known parameters quoted in Theorem 3.4, such that for every $t, s \in [0,T[$, $t < s$, and for every $x \in \mathbb{R}^d$:*

$$|\nabla_x u(t,x) - \nabla_x u(s,x)| \le C_{7.2}(T-s)^{-(1-\beta_{7.2})/2}(s-t)^{\beta_{7.2}/2}.$$

Proof. We assume w.l.o.g that $x = 0$. Note from Theorem 3.3 that we can also assume that $s - t \le T - s$.

From (7.10), for $i \in \{1, \dots, d\}$, $\partial\tilde{u}/\partial x_i(t,0)$, writes (replace the boundary condition $\tilde{G}$ at time T by the boundary condition $\tilde{u}(s,\cdot)$ at time s):

$$\partial_{x_i}\tilde{u}(t,0) = \int_{\mathbb{R}^d} \tilde{u}(s,z)\frac{\partial\psi^{(c)}}{\partial x_i}(t,0;s,z)dz + \int_t^s \int_{\mathbb{R}^d} \tilde{g}(r,z)\frac{\partial\psi^{(c)}}{\partial x_i}(t,0;r,z)drdz.$$

Note from the definition of $\psi^{(c)}$ (*cf.* Theorem 5.1) that $\dfrac{\partial\psi^{(c)}}{\partial x_i}(t,0;s,z) = -\dfrac{\partial\psi^{(c)}}{\partial z_i}(t,0;s,z)$. Thus, deduce from an integration by parts that:

$$\frac{\partial\tilde{u}}{\partial x_i}(t,0) = \int_{\mathbb{R}^d} \frac{\partial\tilde{u}}{\partial x_i}(s,z)\psi^{(c)}(t,0;s,z)dz + \int_t^s \int_{\mathbb{R}^d} \tilde{g}(r,z)\frac{\partial\psi^{(c)}}{\partial x_i}(t,0;r,z)drdz.$$

Thus,

$$\begin{aligned}\frac{\partial\tilde{u}}{\partial x_i}(t,0) - \frac{\partial\tilde{u}}{\partial x_i}(s,0) &= \int_{\mathbb{R}^d} \big[\frac{\partial\tilde{u}}{\partial x_i}(s,z) - \frac{\partial\tilde{u}}{\partial x_i}(s,0)\big]\psi^{(c)}(t,0;s,z)dz \\ &\quad + \int_t^s \int_{\mathbb{R}^d} \tilde{g}(r,z)\frac{\partial\psi^{(c)}}{\partial x_i}(t,0;r,z)drdz \\ &\equiv S(1) + S(2).\end{aligned} \tag{7.16}$$

Derive from Lemma 7.1 that:

$$\begin{aligned}|S(1)| &\le C(T-s)^{-(1-\gamma)/2}\int_{\mathbb{R}^d} |z|^{\gamma}\psi^{(c)}(t,0;s,z)dz \\ &\le C(T-s)^{-(1-\gamma)/2}(s-t)^{\gamma/2}.\end{aligned} \tag{7.17}$$

Deal now with $S(2)$. Following (7.4) and (7.6) and applying Lemma 5.4, we claim for a small $\varepsilon > 0$:

$$\begin{aligned}|S(2)| &\le C\int_t^s (r-t)^{-1/2}\Big[\int_{\mathbb{R}^d} \big|\tilde{g}\big(r, \Gamma^{1/2}(t,r)z\big)\big||z|\exp\big(-\frac{|z|^2}{2}\big)dz\Big]dr \\ &\le C(\varepsilon)\int_t^s \Big[(r-t)^{-(1+\varepsilon)/2}\Big[\int_{B(0,2)} |\tilde{g}(r,z)|^{d/\varepsilon}dz\Big]^{\varepsilon/d}\Big]dr \\ &\le C(\varepsilon)\Big[\int_t^s (r-t)^{-d(1+\varepsilon)/(2(d-\varepsilon))}(T-r)^{-d(1-\gamma)/(d-\varepsilon)}dr\Big]^{1-\varepsilon/d} \\ &\quad \times \Big[\int_t^s \int_{B(0,2)} (T-r)^{d(1-\gamma)/\varepsilon}|\tilde{g}(r,z)|^{d/\varepsilon}dz\,dr\Big]^{\varepsilon/d}.\end{aligned}$$

Deduce now from Theorem 3.5 that:

$$
\begin{aligned}
|S(2)| &\leq C(\varepsilon)(T-s)^{-(1-\gamma)} \left[\int_t^s (r-t)^{-d(1+\varepsilon)/(2(d-\varepsilon))} dr \right]^{1-\varepsilon/d} \\
&\quad \times \left[\int_t^T \int_{B(0,2)} (T-r)^{d(1-\gamma)/\varepsilon} |\tilde{g}(r,z)|^{d/\varepsilon} dz\, dr \right]^{\varepsilon/d} \\
&\leq C(\varepsilon)(s-t)^{1-\varepsilon/d-(1+\varepsilon)/2}(T-s)^{-(1-\gamma)}(T-t)^{\varepsilon/d}
\end{aligned}
\tag{7.18}
$$

Since $s-t \leq T-s$, deduce that $T-t \leq 2(T-s)$ and thus, from (7.18), that:

$$
\begin{aligned}
|S(2)| &\leq C(\varepsilon)(s-t)^{\varepsilon/2}(s-t)^{1/2-\varepsilon/d-\varepsilon}(T-s)^{-(1-\gamma)}(T-s)^{\varepsilon/d} \\
&\leq C(\varepsilon)(s-t)^{\varepsilon/2}(T-s)^{-1/2-\varepsilon+\gamma}.
\end{aligned}
\tag{7.19}
$$

As $\varepsilon \to 0$, the exponent of $T-s$ in (7.19) tends to $-1/2+\gamma$. Thus, for ε small enough and for a new $\gamma > 0$:

$$
|S(2)| \leq C(s-t)^{\gamma/2}(T-s)^{-(1-\gamma)/2}. \tag{7.20}
$$

From (7.16), (7.17) and (7.20), we complete the proof. □

8 Conclusion

As a conclusion, we discuss the strong solvability of the FBSDE (E) and gives further applications of the estimates given in Section 3.

Strong Solvability of (E)***.*** It is rather clear that the FBSDE (E) is not strongly solvable since the forward equation reduces in the decoupled case (*i.e.* $b = b(t,x)$ and $\sigma = \sigma(t,x)$) to a SDE with Hölder continuous coefficients. It is then well-known that this one may not be solvable in the Itô sense: see for instance Barlow [3] for a suitable example.

However, if the coefficient σ is assumed to be continuous in (t,x) and Lipschitz continuous with respect to the variable x, the SDE (4.1) turns out to be strongly solvable (see e.g. Veretennikov [40]): then, the solution built in Subsection 4.2 is also strong. In this frame, the method given in Subsection 4.3 still applies and permits to prove that uniqueness also holds in the pathwise sense.

Decoupling Strategy. This remark emphasizes, to our own point of view, the deep power of the "decoupling strategy" introduced in the earlier *four step scheme* of Ma, Protter and Yong [30]. Indeed, it both applies to the strong setting and to the weak point of view. In short, the crucial point just consists in establishing a suitable integrability property of the singularities of the derivatives of u of order one and two in x in the neighbourhood of the final bound T.

Possible Extensions. Two possible extensions of this work are conceivable. Mention first that our result extends to the multi-dimensional setting provided that the growth of f be at most linear in z (*cf.* to the Heinz example in Krylov [24], page 197, for a counter-example to the unique solvability of a system of PDEs with a quadratic coefficient f). In this frame, the components of the process Y are taken in $S^2_{t,T}(\Omega, \{\mathcal{F}\}, \mathbb{P}, \mathbb{R}^q)$, q standing for the dimension of the backward component. Note however that the proof differs from the usual multi-dimensional strategy: due to the Girsanov procedure in Subsection 4.3, the exponential transform Φ is still necessary to establish the weak uniqueness property.

In the same way, there is no difficulty to weaken the Lipschitz property of f in y to a monotonicity assumption as usually considered in the theory of BSDEs. In particular, according to our previous discussion on strong solutions to (E), we are able to recover the earlier result proved in Delarue [7] with the so-called "induction method".

Note also that the bounds of the second order derivatives of order two can be strengthened if the coefficients a, b and f are Hölder continuous with respect to the variables t and x. In this case, the solution u to the PDE $(\mathcal{E})$ belongs to $\mathcal{C}^{1,2}([0,T[\times\mathbb{R}^d,\mathbb{R})$ and the pertubation strategy used to establish Theorem 5.5 applies and provides the following pointwise estimate of $\nabla^2_{x,x}u$:

$$\forall (t,x) \in [0,T[\times\mathbb{R}^d,\ |\nabla^2_{x,x}u(t,x)| \le C(T-t)^{-1+\alpha_6}, \tag{8.1}$$

for a suitable $\alpha_6 > 0$. We refer to Guatteri and Lunardi [16] for a similar control under stronger assumptions on the coefficients.

Of course, such a bound as (8.1) would simplify in a very sensible way the proof of the unique solvability of (E): taking into account the integrability of the bound in the neighbourdhood of T, a classical Gronwall argument would apply without any difficulties.

For this reason, we feel that the estimate (8.1) could be applied to different asymptotic problems involving the FBSDE representation. Recall indeed from Subsection 2.4 that the "decoupling strategy" permits to deal with homogenization and numerical approximation. We then guess that these works could be extended to the Hölder framework.

References

[1] Antonelli, F. (1993): Backward-forward stochastic differential equations. Annals of Applied Probability, **3**, pp. 777–793.

[2] Antonelli, F., Ma, J.(2003): Weak solution of Forward-Backward SDE's. Stochatsic Analysis and Application, **21**, n.3, pp. 493–514.

[3] Barlow, M. T. (1982). One-dimensional stochastic differential equations with no strong solution. J. London Math. Soc., **26**, pp. 335–347.

[4] Bass, R. F. (1998): Diffusions and elliptic operators. Probability and its Applications. Springer-Verlag, New York.

[5] Brossard, J. (2003). Deux notions équivalentes d'unicité en loi pour les équations différentielles stochastiques. (French. English, French summary). Séminaire de Probabilités XXXVII, 246–250, Lecture Notes in Math., 1832, Springer, Berlin, 2003.

[6] Buckdahn, R., Hu, Y. (1998): Probabilistic approach to homogenizations of systems of quasi-linear parabolic PDEs with periodic structures. Nonlinear Anal., **32**, pp. 609–619.

[7] Delarue, F. (2002): On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. Stochastic Process. Appl., **99**, pp. 209–286.

[8] Delarue, F. (2003): Estimates of the Solutions of a System of Quasi-Linear PDEs. A Probabilistic Scheme. Séminaire de Probabilités XXXVII, pp. 290–332, Lecture Notes in Math., 1832, Springer, Berlin, 2003.

[9] Delarue, F. (2004): Auxiliary SDEs for homogenization of quasi-linear PDEs with periodic coefficients. Annals of Probability, **32**, pp. 2305–2361.

[10] Delarue, F., Menozzi, S. (2004): A Forward-Backward Stochastic Algorithm for Quasi-Linear PDEs. To appear in Ann. Appl. Prob.

[11] Friedman, A. (1964): Partial differential equations of parabolic type. Prentice-Hall, Inc., Englewood Cliffs.

[12] Friedman, A. (1969): Partial differential equations. Holt, Rinehart and Winston, Inc., New York-Montreal, Que-London.

[13] Friedman, A. (1975): Stochastic differential equations and applications. Vol. 1. Probability and Mathematical Statistics, Vol. 28. Academic Press New York-London.

[14] Fuhrman, M., Tessitore, G. (2002): The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. Stoch. Stoch. Rep., **74**, 429–464.

[15] Gilbarg, D., Trudinger, N.S. (1983): Elliptic partial differential equations of second order, Second edition. Grundlehren der Mathematischen Wissenschaften, **224**, Springer-Verlag, Berlin.

[16] Guatteri, G., Lunardi, A. (2005): Smoothing of quasilinear parabolic operators and applications to forward-backward stochastic systems. Advances in Differential Equations, **10**, pp. 65–88.

[17] Hu, Y., Peng, S. (1995): Solution of forward-backward stochastic differential equations. Probab. Theory Rel. Fields, **103**, pp. 273–283.

[18] Hu, Y., Yong, J. (2000): Forward-backward stochastic differential equations with nonsmooth coefficients. Stochatic Process. Appl., **87**, pp. 93–106.

[19] Karatzas, I., Shreve, S.E. (1991): Brownian Motion and Stochastic Calculus, Graduate Texts in Mathematics, 113. Springer-Verlag, New York, xxiv+470 pp.

[20] Kobylanski, M. (2000): Backward stochastic differential equations and partial differential equations with quadratic growth. Annals of Probability, **28**, no. 2, pp. 558–602.

[21] Krylov, N.V. (1979). Controlled Diffusion Processes. Springer-Verlag. New-York.

[22] Krylov, N.V., Safonov, M.V. (1979): An estimate for the probability of a diffusion process hitting a set of positive measure, (Russian). Dokl. Akad. Nauk SSSR, **245**, pp. 18–20.

[23] Krylov, N.V., Safonov, M.V. (1980): A property of the solutions of parabolic equations with measurable coefficients, (Russian). Izv. Akad. Nauk SSSR Ser. Mat., **44**, pp. 161–175.

[24] Krylov, N.V. (1987): Nonlinear elliptic and parabolic equations of the second order. Mathematics and its Applications (Soviet Series), D. Reidel Publishing Co., Dordrecht.

[25] Krylov, N. V. (1996): Lectures on elliptic and parabolic equations in Hölder spaces. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI.

[26] Ladyzhenskaya, O.A., Solonnikov, V.A., Ural'tseva, N.N. (1968): Linear and quasi-linear equations of parabolic type. Translation of Mathematical Monographs, American Mathematical Society, Providence, RI.

[27] Lejay, A. (2004): A probabilistic representation of the solution of some quasi-linear PDE with a divergence form operator. Application to existence of weak solutions of FBSDE. Stochatic Process. Appl., **110**, pp. 145–176.

[28] Lunardi, A. (1995): Analytic semigroups and optimal regularity in parabolic problems. Progress in Nonlinear Differential Equations and their Applications,16.Birkhäuser Verlag, Basel.

[29] Ma, J., Yong, J. (1999): Forward-Backward Stochastic Differential Equations and their applications. LNM-1702, Springer-Verlag.

[30] Ma, J., Protter, P., Yong, J. (1994): Solving forward-backward stochastic differential equations explicitly - a four step scheme. Probab. Theory Rel. Fields, **98**, pp. 339–359.

[31] Nirenberg, L. (1959): On elliptic partial differential equations. Ann. Scuola Norm. Sup. Pisa, **13**, pp. 115–162.

[32] Pardoux, É., Peng, S., (1990): Adapted solution of a backward stochastic differential equation. Systems Control Lett., **14**, pp. 55–61.

[33] Pardoux, É, Peng, S. (1992): Backward stochastic differential equations and quasi-linear parabolic partial differential equations. In: Rzosvskii, B.L., Sowers, R.S. (Eds). Stochastic partial differential equations and their applications, Lect. Notes in Control & Info. Sci, Springer, Berlin, Heidelberg, New York, **176**, pp. 200–217.

[34] Pardoux, É., Tang, S. (1999): Forward-Backward stochastic differential equations and quasilinear parabolic PDEs. Probab. Theory Rel. Fields, **114**, pp. 123–150.

[35] Peng, S., Wu, Z. (1999): Fully coupled forward-backward stochastic differential equations and applications to optimal control. SIAM J. Control Optim., **37**, no. 3, pp. 825–843.

[36] Rogers, L.C.G., Williams, D., (1987): Diffusions, Markov Processes and Martingales : Itô Calculus. John Wiley and Sons Ltd, Chichester, UK.

[37] Stroock, D.W., Varadhan S.R.S, 1(979): Multidimensional diffusion processes. Springer-Verlag, New-York.

[38] Thalmaier, A. (1997): On the differentiation of heat semigroups and Poisson integrals, Stochastics Stochastics Rep., **61**, pp. 297–321.

[39] Thalmaier, A., Wang, F.Y. (1998): Gradient estimates for harmonic functions on regular domains in Riemannian manifolds, J. Funct. Anal., **155**, pp. 109–124.

[40] Veretennikov, A. Yu. (1980): Strong solutions and explicit formulas for solutions of stochastic integral equations. (Russian) Mat. Sb. (N.S.) **111**(**153**).

[41] Yong, J. (1997): Finding adapted solutions of forward-backward stochastic diffrential equations - method of continuation. Probablity Theory Related Fields, **107**, pp. 537–572.